DOCUMENT RESUME

ED 302 042 EC 211 839

AUTHOR Deno, Stanley L.
TITLE Experimental Teaching: An Approach to Improving Student Achievement, Changing Teacher Beliefs, and Identifying Effective Practices.
INSTITUTION Minnesota Univ., Minneapolis.
SPONS AGENCY Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE 85
GRANT G008400649
NOTE 104p.
PUB TYPE Reports - Research/Technical (143) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Disabilities; Elementary Secondary Education; *Experimental Teaching; *Individualized Instruction; *Inservice Teacher Education; *Instructional Effectiveness; Instructional Innovation; Intervention; Outcomes of Education; Performance Factors; Research Methodology; Resource Room Programs; Student Evaluation; Teacher Attitudes; *Teaching Methods
IDENTIFIERS *Curriculum Based Assessment

ABSTRACT

This project sought to test the feasibility and effectiveness of experimental teaching in resource programs, by determining whether teachers could use continuous data to test successive hypotheses and empirically determine which reforms in an individual student's educational program produce the best performance outcomes. The project also evaluated whether learning and implementing experimental teaching procedures had any effect on teachers' knowledge, attitudes, and beliefs regarding individually tailored instruction. Fifty-five special education teachers were taught curriculum-based measurement procedures, time series research methods, and intervention strategies. Analysis of 2 years of research data showed that the teachers were successfully trained to identify effective instructional programs for their students. Student data indicated greater performance increases under a treatment-oriented strategy, where frequent program modifications were made on the basis of direct comparison of progress in successive intervention phases, rather than a more traditional goal-oriented strategy. Results from the measurement of teacher knowledge and beliefs indicated that teachers modified their beliefs concerning the effectiveness of educational interventions. Appendices include a listing and description of instructional contrasts generated by participant teachers, an intervention knowledge scale, and other items. (JDD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302042

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Final Report for Grant No. G008400649

Experimental Teaching: An Approach to Improving Student Achievement, Changing Teacher Beliefs, and Identifying Effective Practices

Stanley L. Deno
Principal Investigator
University of Minnesota

EC 211839

BEST COPY AVAILABLE

Project Objectives

The overall purpose of this project has been to test the feasibility and effectiveness of experimental teaching in resource programs. That is, we have sought to give teachers the tools and training to, in effect, conduct successive experimental tests of the hypotheses generated initially in developing a child's IEP, and subsequently as a child receives special education. This approach emphasized that teaching, especially of handicapped learners, is better viewed as applied research, rather than established technology.

The primary question addressed in this research is whether teachers could in fact use continuous data to test successive hypotheses and empirically determine which reforms in an individual student's educational program produce the best performance outcomes. Teachers participating in the project were taught curriculum-based measurement (CBM) procedures which provided the basis for monitoring student progress. At the same time, teachers were instructed in a variety of program options from which they chose particular interventions which would represent program changes for the students in their classroom. Intervention strategy options and the design for changes in strategies differed from Year One to Year Two of the study. During Year One, teachers tested whether their original program plans or one of two specific contrasted intervention changes was most effective for their individual students. During Year Two, the original program plans were contrasted with one major teaching strategy which was supplemented with additional instructional components according to one of two CBM decision rules.

A second major issue addressed by this research is whether learning and implementing experimental teaching procedures would have any effect on

teachers' knowledge, attitudes and beliefs regarding individually tailored instruction. A major assumption of this research was that no one instructional strategy would be effective for all students. We believed that as teachers systematically introduced new instructional approaches and tested those approaches with frequent CBM procedures, the individual differences in intervention effectiveness would be evident in students' measured performance. We expected that repetitions of this approach would alter teachers' beliefs about the need for individually tailored educational programs.

Year 1 Methodology

Administrative

Project staff were identified and assigned to participate in the project starting July, 1984. The Assistant to the Director position was split between two doctoral students in Educational Psychology, Ann Casey and Russell Skiba. In addition, Douglas Marston, a psychologist in the Minneapolis Public Schools was hired as the Minneapolis liaison for the project. Deanne Magnuson with the Minneapolis Public Schools' Department of Special Education provided invaluable administrative support throughout every phase of the project. Six resource specialists from this department were recruited as teacher trainers for the project and worked closely with project staff throughout the implementation phases of the study.

Subjects

Subjects volunteering for the study were 28 special Education Resource Teachers (SERTs) currently teaching in the Minneapolis K-6 School Based Resource Program. These teachers had extensive backgrounds in education.

Eighty-two percent of the teachers were certified to teach in regular education. In special education certification areas, 89% of the SERTs were certified in learning disabilities (LD), 18% in emotional disturbance (ED), 50% in educable mental retardation (EMR) and 7% in trainable mental retardation (TMR). Thirty-nine percent of the SERTs had special education certification in at least 2 areas. The average length of time SERTs had taught was 12.6 years; 10.2 years in special education and 2.4 years in regular education.

For the study each SERT worked with 4 mildly handicapped, elementary grade level students. The distribution of students in each grade level was as follows: 4 pupils at 1st grade; 11 at 2nd grade; 14 at 3rd grade; 27 at 4th grade; 22 at 5th grade; 29 at sixth grade; 1 at 7th grade; and 3 at 8th grade. The average grade level of the 111 pupils was 4.42. The reading skills of each student were at or below 50% of their grade level norm.

Training Procedures

The original proposal outlined the necessity for training teachers in three areas: curriculum-based measurement, time series research methods, and program alternatives. The first two areas were required to provide teachers with a methodology for testing the efficacy of various program alternatives, the third for providing training in a specific set of programming variables to implement and test throughout the course of the research project. As curriculum-based measures (CBM) were previously developed and tested in research programs at the University of Minnesota, these required only adaptation for project purposes. An experimental teaching methodology unit was developed by project staff in July, 1984.

Other program alternative units included instructional variables, motivation variables, curriculum variables, consultation variables, cognitive strategy variables, and peer tutoring variables.

In early Fall of 1984, teacher trainers met with teachers to provide instruction in the CBM system. Although most teachers had some previous knowledge of the system, further training was necessary to ensure standard implementation of the measurement system. In addition, both experimental teachers and 2 groups of control teachers completed three pre-tests designed to measure knowledge, beliefs and attitudes toward instructional interventions. The first control group consisted of 20 K-6 resource teachers in the Minneapolis school district who agreed to participate in the project during Year Two of the study. The second control group consisted of the remaining 33 K-6 resource teachers in the district who did not volunteer for the project. The control groups were comparable to the experimental group in experience with an average of 16.2 years of teaching experience.

In November, a week was set aside for training teachers in experimental teaching and programmatic interventions. To avoid conflict with regular teaching schedules, and to maximize the perception of professional growth for such training, substitutes were hired for teachers for the entire week and training sessions were held in conference centers away from the schools. The first four days of training consisted of the presentation of the program alternative modules and the research design module. The fifth day was set aside for teachers to consider which interventions they planned to implement. They were asked to generate intervention contrasts,

interesting to them, that would constitute a major change in an instructional program for reading. In total, the special education teachers listed 73 instructional contrasts, which may be found in Appendix A.

One additional day was set aside three weeks later for teachers to make final decisions regarding which interventions they planned to implement. Over the course of the day, participants formed themselves into groups of at least three teachers who agreed to test the same two contrasting experimental treatments. On this basis the following 8 contrast groups were formed.

Group	Treatment #1	Treatment #2
1	Whole Word in Holt	Phonics in Holt
2	Whole Word in Holt	Merrill Linguistic
3	Language Experience	Oral Reading & Token System
4	High Interest, Low Vocabulary with Dolch Sight Words	Language Experience
5	Phonics with Cross-Age Tutor	Free Reading with Cross-Age Tutor
6	Individual Reinforcement	Group Reinforcement
7	Precision Teaching without Rate Building	Precision Teaching with Tokens
8	Free Choice Reading Activities	Language Experience

Groups then wrote the specific procedures that would form the 'B' (January 15 to March 15) and 'C' (March 15 to May 15) phases of the experiment (the 'A' phase being the prior baseline period from October 31 to January 15). A detailed description of each instructional strategy may be found in Appendix B. Finally, teachers transferred these descriptions

to revisions of their Individual Objective Plan (IOP) for each student. Each plan included specification of instructional procedures, group arrangement, time allotted for activities, instructional material and motivational strategies employed. The completed worksheet used for the documentation (an example is given in Appendix C) was checked closely by the resource specialists serving as teacher trainers.

To further ensure that each group's instructional contrast was consistently implemented, each teacher was observed twice during Phases B and C by trained observers. Finally, each group met for a half day on four occasions during the course of the project. At these meetings teachers compared notes and discussed the progress of their instructional plans and adherence to standardized measurement and intervention procedures.

CBM Measurement Procedures

The reading achievement of pupils participating in the project was determined with the curriculum-based measurement of reading (Deno, Mirkin, & Chiang, 1982). Following these procedures, students were asked to read aloud for one minute periods from the district's mainstream reading series, the Holt Reading Series. At the end of each reading sample the teachers tabulated the number of words read correctly and charted the data on equal interval graph paper. For each pupil the reading material corresponded to the reading level in which the IEP goal had been set in Holt. All students were presented with randomly drawn passages from the appropriate Holt level three times each week for the entire academic year. Pupil progress within each experimental phase was then determined by calculating the slope of improvement (i.e. the average increase or decrease in number of words read

correctly per week) for each child.

Teacher Knowledge and Attitude Measures

Three measures were developed in order to assess teacher knowledge and attitudes related to individualized instructional interventions. First, a multiple choice knowledge test was developed to test teachers' knowledge of the material presented in the week-long training workshop. This material consisted of the six program alternatives and the experimental teaching design. One could hypothesize that a difference in beliefs about intervention effectiveness might be due to differences in knowledge about various program alternatives. Thus, both experimental and control groups were administered this knowledge test, prior to training, to determine if there were any pre-existing group differences.

The second measure was an Intervention Knowledge Scale (see Appendix D). Two brief case studies were given to the teachers and they were asked to generate as many intervention strategies as they could within a four minute time period. After they completed this task, the teachers were asked to rate the effectiveness of each strategy they had generated using a five point Likert scale. In developing the measure, test-retest reliability data was obtained with a Pearson Product-Moment Correlation coefficient of .683 ($p = .001$) for number of strategies and .969 ($p = .001$) for the effectiveness rating.

The third measure developed was the Belief Scale (see Appendix E). Teachers were given a list of 30 teaching strategies and were asked to estimate the percentage of students for whom each strategy would be effective in increasing a student's reading achievement over a six week

period. They were also asked to rate the certainty on which these judgments were held using a one to five Likert scale. Test-retest reliability of the measure was .888 ($p = .001$) for the effectiveness estimates and .888 ($p = .001$) for the certainty rating.

Teacher Judgment of Intervention Effectiveness

At the end of phase C, the experimental group of teachers participated in a workshop on techniques to aid in the visual analysis of time-series data. The workshop consisted of three parts. In part one, teachers were asked to complete a questionnaire that required them to make decisions about each student's graphed data. For each student, the teacher was asked whether reading performance increased more in the B phase, in the C phase, or whether the two phases represented equal improvement. Teachers were also asked to rate the certainty of their answer on a five-point scale, and to list the criteria (level, slope, variability, or other factors) they used in making their decision. Each student's graph was rated by both the teacher responsible for that student and another teacher, who served as the reliability observer.

In part two, teachers were trained in a technique to enhance the effectiveness of visual analysis known as the quarter-intersect method for analyzing slope (White & Haring, 1980). Quarter-intersect slopes were calculated for each teacher's four students and the slopes were checked by the reliability observer, as well as by teacher trainers. Slopes on which there was less than 80% agreement between the two teachers were recalculated until the criterion was met.

Finally, in part three, the participating teachers were again asked to complete a questionnaire identical to the one completed during the first

part of training. Both the teacher and the original reliability observer completed the questionnaire for each of the four students measured.

Year 1 Results and Discussion

Teacher Knowledge and Attitudes

The first teacher measure on knowledge of program alternatives was administered in order to control for teacher differences before training and to determine how much the experimental teachers had learned from the training provided. On the pretest administration, the mean number correct out of 30 questions was 16.48 for the experimental group, 16.88 for control group 1, and 15.89 for control group 2 resulting in a non-significant F-test for differences among the groups. Thus no pre-treatment differences in intervention knowledge were evidenced. On the second administration of this instrument, after teacher training was complete, the experimental group increased their performance (mean correct = 22.50) and was significantly different from the two control groups (with means of 17.75 and 15.26 respectively). Thus it was found, as expected, that teacher training did result in a positive change in overall knowledge related to alternative interventions.

An additional measure of knowledge related to interventions was provided by the Intervention Knowledge Scale which yielded data on the number of relevant strategies which teachers could list within a four minute period of time when presented with a short case study. At pretest administration (i.e. time 1) the number of strategies generated by the experimental and control 1 groups were very similar (with means of 17.00 and 17.62) while control group 2 listed significantly fewer strategies (p =

.014). Given the fact that the teachers were not randomly assigned to groups and chose to participate in the project, this is not a surprising result. It may be that teachers who have more intervention alternatives available to them are also the same as those who chose to participate in the project.

Insert Table 1 about here

The consistent pattern of means across times 2 and 3, reflected in Table 1, indicates that the second control group maintained its relative standing with significantly fewer interventions generated ($p < .001$) even as all 3 groups significantly increased the number of interventions they could list ($p < .001$) with repeated administration of the instrument.

Table 2 presents the data on the percent of students estimated to be positively effected by the interventions listed on the Intervention Knowledge Scale.

Insert Table 2 about here

The ratings of effectiveness of the previously generated intervention strategies produced a significant ($p < .04$) group by time interaction, in a repeated measures analysis of variance, with the experimental group rating strategies as less effective over time. At the post intervention (time 3), the experimental group rated strategies with a mean effectiveness rating of 3.48; whereas the two control groups rated strategies as more effective

with mean ratings of 3.75 and 4.01. This interaction is graphically depicted in Figure 1 below.

Insert Figure 1 about here

A second major research question dealt with teachers beliefs regarding the need to individualize students' programs. The Belief Scale was developed for this purpose and two pieces of data were collected. The initial hypothesis was that experimental teachers would decrease their estimates of the percent of students for whom particular strategies would be effective, and that they would become less certain of their answers. Table 3 shows that a significant ($p < .001$) group by time interaction was found in a repeated measures analysis of variance for mean effectiveness percentages.

Insert Table 3 about here

This interaction reflects a pattern of means where experimental teachers lowered their estimates of the proportion of students for whom specific intervention strategies would be effective across time (i.e. from 62.45% at time 1 to 49.82% at time 3), while control teachers remained basically constant in their estimates, (with means ranging from 61% to 68% over the three testing periods). This relationship is graphically depicted in Figure 2 below.

Insert Figure 2 about here

The fact that decreases in effectiveness estimates for the experimental group did not take place until the spring post-testing was expected since training in time series slope analysis and knowledge of the aggregated student outcomes was not presented until the final training session in May. However, the predicted decrease in certainty ratings at time three was not found. The statistically significant main effect difference, the repeated measures ANOVA shown in Table 4, reflects an unexpected change from time one to time two which was consistent across groups. No interaction effect was found for this measure.

Insert Table 4 about here

The focus of the dependent measures developed for this project was on teachers' knowledge and beliefs related to various intervention strategies for reading. The Intervention Knowledge Scale measured the number of strategies teachers could generate in response to a specific case as well as their rating of effectiveness for each strategy. Although there appeared to be a testing effect, since all three groups generated more strategies over time, the experimental group generated the most strategies. In the rating of effectiveness of these strategies, it was hypothesized that experimental teachers would become more moderate in their ratings as a function of participation in the project. This hypothesis was born out as

effectiveness ratings over time showed differential patterns for experimental and control teachers. Teachers who, at the end of the project could identify more intervention alternatives, were more conservative in their estimate of how many students would be positively effected by their suggested strategies. Teachers who indicated interest in the project, but did not participate until Year Two, did not demonstrate such attitude change nor did the controls who showed no interest in participation.

It was also hypothesized that experimental teachers would become more moderate in their beliefs about the percentage of students for whom particular intervention strategies would be effective. It was the researcher's expectation that after participation in the project, teachers would come to believe that some strategies would be effective for certain students and not for others. The results again proved this hypothesis to be true. Interestingly, it appears that inservice training alone did not produce this effect. It was only after both intervention phases were completed and the knowledge of results shared and discussed that changes in belief occurred. This finding is consistent with social psychological research indicating that direct behavioral experience with attitude objects produce the largest and most consistent changes in attitude and subsequent behavior. In other words, it was the <u>experience</u> of experimental teaching which appears to have effected changes in teacher belief related to individualized instruction.

The knowledge of results of continuous time series measures of reading performance for individual students and of the aggregate results of slopes of performance for groups of students treated with contrasted "high impact"

interventions is considered to be the major independent variable operative in the experimental teaching effect. A summary of these results follows.

Student Outcomes under Contrasted Treatments

Student achievement data was analyzed statistically by regression procedures using a Fortran program designed specifically for the experimental teaching data set. Using these procedures, the slope of progress on number of words read correctly from one minute reading samples was calculated for the baseline period (consisting of 8 to 10 weeks) and both intervention phases (each consisting of 8 weeks) for each individual student.

The number of data points on which each slope was based varied somewhat from student to student due to absenteeism, change of teachers, etc. It was decided that only slopes based on at least 12 data points per phase would be included for individual and group analyses. With the exclusion of subjects with insufficient data for one or more of the three phases, the number of subjects included for analysis dropped from 111 to 95.

The calculated slopes of progress represent increases or decreases in the number of words read correctly per calendar week. Thus, if student performance was not measured over a two week period in the middle of the phase, (e.g. due to illness or vacation), the best fit regression line represents gain over the full 8 week phase, rather than the 6 weeks during which measurements were taken.

Insert Figure 3 about here

Figure 3 presents the mean slopes of improvement during each phase of the experiment across all students. The mean levels of performance at the beginning and end of each phase, as well as the actual mean slopes for each phase, are reflected in this figure. Since the metric employed has not been converted to any form of derived or norm based score, the actual increase in reading performance for this group of special education students is accurately represented. We feel that this curriculum based metric (i.e. in raw score terms) provides useful descriptive as well as comparative data for assessing performance across time.

The mean slope for baseline (phase A) was .83 words per week, 1.41 words per week for phase B interventions, and 1.04 words per week for phase C interventions. The mean differences among phases was found to be significant in a repeated measures ANOVA ($F = 4.96$, $p < .01$), with followup t-tests indicating that phase B was significantly different from phase A ($t = 3.11$, $p < .01$) and phase C ($t = 1.99$, $p = .05$). The mean for phase C interventions was not found to be statistically different from phase A.

While it was encouraging to find that in general the experimental interventions during phase B were significantly different from baseline, our main experimental interest for Year One data was in the specific treatment contrasts between interventions within each group of teachers. The breakdown of mean slopes by phase and group is presented in Table 5 and Figure 4 below.

Insert Table 5 about here

Insert Figure 4 about here

A repeated measures analysis of variance was performed in order to determine whether there were differences among the means of the three phases of the experiment within each of the contrast groups. The results of this analysis are presented in the top portion of Table 6. Significant differences were found between phases for both the second and sixth contrast groups. Followup comparisons using pooled variance t-tests were then performed to determine whether phase A (baseline) was significantly different from either of the intervention phases and whether the two intervention phases differed significantly from each other within these two groups. Results of these analyses are presented at the bottom of Table 6.

Insert Table 6 about here

Within contrast group 2, significant differences were found between the mean baseline slope (.559) and the mean slope during phase B intervention (2.171) where the Holt Basal Whole Word curriculum was used during instruction. A significant difference was also found between the mean for phase B and the mean for phase C (.749) where the Merrill Linguistic curriculum was utilized for instruction. The difference between phase C and baseline was found to be non-significant.

Within contrast group 6, significant differences were found between the baseline mean (.473) and both the mean for phase B (1.816), where an

individual reinforcement strategy was used, and the mean for phase C (1.573), where a group reinforcement strategy was used. No difference was found between the group and individual reinforcement strategies.

Fairly substantial variance between subjects (see Appendix F for the range and standard deviations of slopes) within contrast groups limited the number of significant differences found in this analysis. Concern for the equality of variance covariance matrices led to a separate analysis of phase differences within groups. This two-way ANOVA (subjects by phases) yielded essentially the same results, with one exception. Within contrast group 7, the precision teaching group, where components of rate building were withdrawn during phase B and included along with token reinforcement in phase C, there was a significant difference among phases ($F = 6.02$, $p < .01$) and significant followup t-tests for phase C vs. phase A ($t = 3.44$, $p < .01$) and for phase C vs. phase B ($t = 2.12$, $p = .04$).

Followup tests were also conducted to determine whether there were differences among the means for intervention strategies within each of three phases. These tests were conducted in two stages. First an overall comparison among each of the eight group means was performed and then, if the overall test was significant, pair wise comparisons were performed. T-tests were performed using Kirk's (1968, p. 265) procedure for pooling error terms and degrees of freedom (from Table 2) when expected mean squares of the errors were not identical. To control for inflated Type 1 errors, the Bonferroni comparisons were performed to determine critical values. Results of these analyses are presented in Table 7.

Insert Table 7 about here

The results presented in Table 7 indicate that no significant differences were found among the contrast groups in either phase A or phase C. The differences among the eight intervention groups in phase B were found to be significant. Followup tests on all possible pairs of means within phase B found that the four groups with the highest means (groups 1, 2, 5 and 6) were significantly different from the mean for the Language Experience group. These results cannot be viewed as the result of inflated type one error since each of these t-tests were run at an alpha level of .0018 (i.e. .05/28). Since the differences within phase A (baseline) were found to be significant, there was no need to assume that the results in phase B were due to preexisting differences between the mean slope levels before the beginning of the experiment. However, an analysis of covariance adjusting for any preexisting differences was performed to corroborate this assumption. The results in this case were found to be virtually identical, with the same groups found to be significantly different.[1]

One final analysis was done on group differences between mean slopes for the three groups which chose to contrast the Language Experience intervention with another treatment for reading. Since the expectation of most teachers participating in these groups was that Language Experience would be superior to the contrasted intervention, it was interesting to note that this hypothesis was disconfirmed. A two way (group by treatment) analysis of variance showed that the mean slope for the 36 students

involved in the Language Experience intervention (.41) was significantly different ($F = 5.76$, $p = .03$) from the mean of the contrasted interventions (1.00). It was also interesting to note that the Language Experience mean slope was less than the baseline mean slope (.79), though this difference was not statistically significant.

The investigators had no specific hypotheses regarding which of the interventions for reading would be found to be superior. The fact that the Language Experience intervention faired less well than other treatments was more of a surprise to the teachers involved in developing and implementing this strategy. Another finding which surprised many of the teachers was the overall difference in favor of whole word approaches to reading instruction over against phonics approaches. On the other hand, teachers were genrally not surprised to find that strategies employing systematic positive reinforcement proved to produce performances superior to baseline and non-reinforcement conditions.

While the results presented above indicate the superiority of some interventions over others when analyzed on the basis of group differences, it is also instructive to view the results on a single subject basis. Prior to intervention, teachers were asked to predict what proportion of the students would show superior performance under each of the contrasted treatments. Teachers defined an important difference in slope to be a difference of 1.25 words per week slope between the phase B and phase C interventions. Table 8 presents the actual percentage of students fairing better under phase B, those doing better under phase C, and those in which no practical difference was found, using the 1.25 difference criterion.

Insert Table 8 about here

The analysis of the proportion of students fairing better under each contrast condition, presented in Table 8, parallels some of the findings of mean differences, for example, between whole word and phonics approaches. However, it also indicates the large proportion of cases in which no substantial difference between treatments was evidenced. In general, teachers estimated a much higher percentage of students who would benefit from one of the two contrasted interventions and a much lower percentage for those showing no difference than was actually the case.

We believe that the outcomes presented here are directly related to the changes in teacher beliefs presented in the prior section. Throughout the year, teachers received continuous feedback on the progress of their students. At the end of each phase they were able to compare the relative effectiveness of their intervention with the previous period. These results may have been contrary to their original expectations in many cases.

Teachers may have found that the intervention they presumed to be most effective was not as effective as another. In other cases they found that neither intervention appeared successful. In general, the results would indicate they found that interventions which worked for one student didn't necessarily work for all students. As a result of this knowledge of results, they tended to decrease their estimates of the effectiveness of interventions they would propose to deal with specific reading problems and

to estimate a lower percentage of cases in which some of the standard approaches to teaching would be successful.

In summary, we would expect teachers who participated in the project to be more willing to experiment in trying different interventions with different students and to rely on the data to determine the effectiveness of their strategies. We found that analysis of curriculum based measures via time series analysis can yield evidence of the superiority of one intervention over another for individual students and groups of students, and that this evidence may be contrary to previously held beliefs. The next section presents results related to the training provided to aid teachers in making judgments related to the effectiveness of their interventions.

Outcomes of Teacher Judgment Training

Three main questions were considered in analyzing the data generated from the pre and posttest measures of teacher judgments regarding intervention effectiveness. The first question concerned the effect of familiarity with the student on whose graph the judgments were made and the effect of the training in the quarter-intersect method on estimates of judgment certainty. Ratings of certainty taken from the questionnaires completed by teachers familiar with the four students assessed and the reliability observer, who was not familiar with the students, were averaged to produce a mean certainty rating for each. A repeated measures analysis of variance was performed with familiarity and presence or absence of judgmental aid (i.e. the quarter-intersect method) as the within subject factors for this design.

A second question concerns the effects of certainty and judgmental aid on the reliability of teacher's decisions regarding graphed data. Graphs for which teachers and reliability observers agreed on which phase (B, C, or same) showed superior progress were counted as agreements. The proportion of graphs on which each teacher-reliability observer pair showed agreement was the dependent variable in a one-between, one-within repeated measures analysis of variance. Four categories of teacher certainty were established, based on certainty rating before and after the introduction of the quarter-intersect method. The four categories were: high certainty before and after introduction of the judgmental aid (high-high); high initial certainty before which decreased after introduction of the judgmental aid (high-low); low initial certainty that increased with use of the judgmental aid (low-high); and low certainty both pre and post training (low-low).

Finally, the third question addressed whether a shift occurred in the criteria used by teachers to make decisions regarding treatment effects as a result of the training in the use of a judgmental aid. Chi-square Goodness-of-Fit tests were used to answer this question. The number of graphs on which teachers indicated that they had used level, slope, variability, or a combination of criteria on the pretest were entered as the obtained frequencies.

The first repeated measures analysis of variance addressed the effect of familiarity with the subject and the introduction of the judgmental aid on the certainty decision made regarding the relative effectiveness of the two interventions. On a five point scale (1 being most certain and 5 least

certain), teachers' certainty regarding their judgment increased from a mean rating of 2.57 prior to training in the quarter-intersect method, to 1.72 after training, while reliability observer means before and after training were 2.81 and 2.23, respectively. Both main effects in the repeated measures ANOVA were significant. Teachers familiar with the student represented by the graph expressed more confidence in their judgments of student progress than did reliability observers unfamiliar with the student ($F = 4.07$, $p < .05$). At the same time, both teachers and reliability observers' ratings of certainty increased significantly as a result of training ($F = 18.66$, $p < .001$). The interaction of familiarity and use of the judgmental aid was not significant.

The second analysis addressed the effects of teacher certainty and the presence or absence of the judgmental aid on the reliability of teacher decisions regarding graphed data. Means from the four levels of certainty both prior to and following training in the use of the quarter-intersect techique are graphed in Figure 5. Results of the repeated measures analysis of variance (Table 9) indicate no significant main effect for the between groups factor, certainty of judgment. However, the use of the quarter-intersect technique greatly improved the level of agreement between the judgments made by teachers and those made by reliability observers, from .56 prior to training, to .78 following training ($F = 10.09$, $p < .01$). The degree of teacher certainty and the use of the judgmental aid also interacted significantly to increase the level of agreement between teachers' and reliability observers' judgments ($F = 9.03$, $p < .01$). As can be noted in Figure 1, this effect was most pronounced for those cases in

which the introduction of the judgmental aid increased teacher certainty, although some increase in agreement can be noted for those teachers who were highly certain of the judgment in both phases.

Insert Figure 5 about here

Insert Table 9 about here

The final analysis concerning the criteria used in judging the relative effectiveness of the two treatments reveals that both teachers and reliability observers used level, slope, variability or a combination of these criteria in roughly equal proportion prior to training (Table 10). Following training in and use of the quarter-intersect method for drawing and analyzing trend lines, both teachers and reliability observers used slope as the sole criteria in their judgment about treatment effect in 70-80% of the graphs that were viewed. Goodness-of-fit tests indicated that the shift in distribution was statistically significant for both the teachers (chi-square = 56.66, $p < .001$) and the reliability observers (chi-square = 59.78, $p < .001$).

Insert Table 10 about here

The results presented in Table 10 suggest that a shift of teachers' use of the data resulted from the introduction of the judgmental aid. The

shift to almost exclusive reliance upon slope as a basis for judgment presents somewhat of a paradox since use of multiple criteria is most often recommended by experts and was in fact the method advised in the teachers' original training in single subject design. Yet, the level of agreement between teachers and reliability observers did improve as they focused their attention on the single characteristic of slope, to the exclusion of level and variability. It may be that by focusing attention on a single characteristic of the data, attention is drawn away from subjective factors such as prior knowledge of the student. This in turn may result in higher levels of agreement with observers who have had no prior contact with the student represented by the graph.

Training in the quarter-intersect technique for fitting a line of progress to the data resulted in an increase in the proportion of graphs on which the teacher and reliability observer agreed in their judgment about treatment effects. The increase in the level of inter-observer agreement, from .56 prior to training to .78 following training, indicated that the addition of such a judgmental aid may be important to ensuring the reliability of teacher judgment concerning graphed data.

This training also resulted in a marked increase in the confidence expressed in teacher judgment of treatment effects. The average certainty rating for both teachers and reliability observers increased significantly after training in and use of the quarter-intersect method.

Prior familiarity with the history of the subject whose performance was graphed was also related to certainty of judgments about treatment effects. Teachers rated themselves as more certain of their judgments about

treatment effects than did reliability observers. The results of the Chi-Square Goodness of Fit tests suggested that this difference was more pronounced prior to training in the use of the judgmental aid.

The relationship between the certainty teachers have in their judgments of treatment effects and the reliability of those judgments is less clear. There was no significant main effect differences in the level of inter-observer agreement for any of the four categories of certainty of judgment. The same analysis did indicate, however, a significant interaction of certainty and use of the judgmental aid in determining the reliability of visual inference. The correlation between teacher certainty and reliability of judgment did not appear until after training in the use of the judgmental aid. After training, the two groups of teachers who were more certain of their judgments also showed higher levels of agreement with reliability observers than did the two groups who were less certain of their judgments regarding treatment effects.

Taken as a whole, these results suggest that the training provided in the quarter-intersect method of drawing progress trend lines was effective in increasing the reliability of teachers' judgments and increasing the certainty of treatment effect judgments based on important characteristics of the data. These encouraging results provided impetus for including training in the quarter-intersect method for estimating slopes of progress and judging the differential effects of treatment interventions as a key element in the training of Year Two experimental teaching participants.

Year 2 Methodology

Administrative

During the second year of the project, the Assistant to the Director position was held by David Heistad, a second year doctoral student in Educational Psychology. School Psychologist, Douglas Marston, again served as Minneapolis liaison for the project. Eight resource specialists from the Minneapolis Public Schools Department of Special Education, six of whom had served as teacher trainers during year 1, played a very active role during year 2 in developing and implementing training modules. Valuable administrative support was again provided by Deanne Magnusson, Keith Kromer, and Judy Brown in the Minneapolis Special Education Department.

Subjects

Subjects volunteering for year 2 of the Experimental Teaching Project are similar in background to year 1 teachers, with an average of 16.2 years of teaching experience for the 27 teachers.

Each teacher worked with 4 mildly handicapped, elementary grade level students over the course of the entire 1985-86 school year. The distribution of students in each grade level was as follows: 3 pupils at 1st grade; 19 at 2nd grade; 26 at 3rd grade; 19 at 4th grade; 20 at 5th grade; 15 at 6th grade; and 1 at 7th grade. The average grade level of the 103 pupils was 3.09. The reading skills of each student were at or below 50% of their grade level norm.

Training Procedures

During the second year of the project, teachers were again trained in the areas of curriculum-based measurement, time series research methods and

program alternatives. However, the format of this training changed considerably from year 1 to year 2.

Early in the fall of 1985, project staff, teacher trainers, and participating special education teachers met for two full days of training. During the first day, project staff introduced the essential features and assumptions on which the experimental teaching paradigm is based and reviewed the project objectives. A review of the measurement system methodology and literature on the reliability and validity of curriculum-based measurement was also presented. Teachers were then given instruction and practice in the quarter-intersect method of drawing and comparing trend lines until competence in this skill was demonstrated by each teacher. During this first day, the teacher measures of intervention knowledge and attitudes related to interventions were administered in the same manner as year 1.

During the second day of the workshop, teachers were presented with introductions to the program alternative modules which would serve as the chief independent variables for year 2 research. The eight training modules were: Direct Instruction; Peer Tutoring; Edmark Reading Program, a packaged computer directed curriculum; Computer Software for Reading, an individualized curriculum using Essential Sight Word Programs (ESWP); SIMS, a task analytic approach to coding; Reading Comprehension and Cognitive Strategies; Precision Teaching; and Contingency Management. Following introduction of the program alternatives, teachers met with teacher trainers to get more specific information on the appropriateness of various alternatives for intervention with their four special education students.

Finally, they decided which of the program alternative modules they would attend.

Each teacher was allotted 4 days of release from class responsibilities to attend training sessions on the program alternatives. Each module had a specified schedule and time commitment (e.g. direct instruction required 2 days instruction, computer software 3 days, and contingency management one half day) which guided teachers in developing their own combination of modules for training. A list of the module and number of days in which teacher participated is given in Appendix G.

Training in the various instructional modules for reading was begun on October 16 and finished by November 14 of 1985. During this period (Phase A) teachers collected baseline data on the curriculum-based measures for reading using the standard procedure outlined for year 1. They were asked not to intervene with any changes in instructional methodology until all training in their modules was complete and specific procedures for their intervention were documented. At the end of the training period, each teacher was asked to choose one of the eight program alternatives to implement with each of their 4 students during phase B of the study. However, two teachers, who worked with more than 4 students, were permitted to participate in two different interventions.

Only 5 of the 8 program interventions were selected by teachers for intervention during phase B. The distribution of teachers by group was as follows: 15 teachers chose Direct Instruction; 4 teachers intervened with Precision Teaching; 1 teacher with SIMS; 1 teacher with Cognitive Strategies; 4 teachers with the Edmark Reading Program; and 4 teachers with

the ESWP/computer program. Once these groups had been determined, teachers documented the specific procedures which constituted the baseline condition for each student and wrote detailed Individualized Objective Plans for the specific intervention changes to be carried out during phase B of the experiment.

Each group had a somewhat different timeline for implementation of their phase B interventions. For the majority of the teachers, the baseline (phase A) condition lasted from late September until the middle of November. However, the Edmark Reading Program group began phase B on October 28 while the Essential Sight Word/Computer intervention group began phase B on November 5. Phase B interventions were conducted until January 17 for all groups.

To insure that each group's instructional intervention was consistently implemented, teachers were observed in their classrooms by teacher trainers, some of whom used videotape for feedback and correction. Also, each group met for a half day on four occasions during the course of the project for "chart sharing" meetings. At these meetings teachers compared notes and discussed the progress of their instructional plans and adherence to standardized measurement and intervention procedures.

Data Presentation and Instructional Evaluation Procedures

During the course of the year participating experimental teachers had numerous opportunities to present their data. This served as an important aspect of training in the evaluation of the specific interventions but also turned out to be an effective stimulus for discussion of program alternatives and overall concerns related to the instruction of special

education students. Following one session of these presentations, at the midyear meeting with all year teachers, the program coordinator for Minneapolis Special Education Resource Programs, Deanne Magnusson and one Elementary School Principal led an open discussion regarding the opportunities and constraints facing special education teachers in Minneapolis. They highly praised the development of teacher expertise, not only in the administration and evaluation of curriculum-based measures, but also of the skill teachers demonstrated in presenting and discussing student progress and concerns from a data based perspective.

On March 12, a "reunion" of teachers from Year 1 of the project was held so that they might receive feedback on the analysis of results on student outcomes and teacher knowledge and belief measures. During this half day program, four of the teachers from the year 2 project presented graphs on the progress of their students. Their presentations were well received and again indicated a level of expertise comparable to what is found at professional conferences.

Teachers from year 1 were also asked to fill out a questionnaire evaluating their training and research experience in the project. Responses to this questionnaire indicated that 83% of the year 1 teachers had initiated at least one new teaching strategy during the subsequent year. The average number of modification made on the basis of graphed data was reported to be 7.05. One of the most frequently mentioned benefits of the project was the establishment of a forum where special education teachers could share their concerns with other teachers and where a sense of professional camaraderie was reinforced.

Data Utilization Procedures

During the second chart sharing meeting in January, the details of the phase C design for the year 2 study were presented to each group. Following a review of the quarter-intersect method for establishing trend lines, each teacher drew in appropriate slope estimates on each of their students' graphs. Each teacher then had an opportunity to present one of their graphs to the rest of the teachers in their group and comment on the progress made and any apparent problems or difficulties experienced with particular students. Teachers were then encouraged to brainstorm possible intervention changes which would supplement the main intervention strategy employed during phase B. For example, a direct instruction (DI) teacher suggested adding a contingency management component to her group instruction by employing tokens along with the verbal praise already standard to the DI approach. After considering what types of intervention changes might be made, teachers were then instructed in two different criteria for when and to what degree a change should be made.

With the exception of the Edmark and ESWP/Computer groups, teachers were randomly assigned to one of two data utilization strategies which would serve as the major contrast during phase C of the study. The first data utilization strategy was called the "treatment oriented" approach. It necessitated that some change had to be made after each period of 9 to 12 one minute reading sample measurements. (That is, no change could take place before 9 data points had been collected, but some change had to be made by the 12th data point.) At the end of phase B data A slope line was drawn using the quarter-intersect method. This trend line thus served as the reference point for comparison to the first period of 9 to 12 data

points in phase C. Teachers were instructed to draw a new slope of improvement for this first time series in phase C (typically 3 to 4 weeks of data), and to decide what type of intervention change needed to be made by comparing this slope with the slope of phase B. If the rate of progress indicated by this slope line was higher than the previous period, minor modifications or additions to the instructional program were advised. If, however, the rate of progress for the 9-12 data point period was less than the previous period more substantial changes could be considered. Thus, the degree of modification advised was relative to the magnitude of the difference in slope between two consecutive periods of time. This procedure of drawing a slope line every 9-12 data points, with an aim towards maximizing the slope in successive periods, was continued throughout the remainder of the year for the "treatment oriented" group.

The second data utilization strategy was called the "goal oriented" approach. The teachers assigned to this group also drew in phase B slope lines for all of their students. They too were instructed to make some change at the beginning of phase C. However, after this point, their basis for deciding when to change differed markedly from the "treatment oriented" approach. They were instructed to locate the point on each graph where the slope line intersected the line indicating the end of phase B and to draw a goal line from this point according to one of two criteria. The rules for deciding which criteria was appropriate were as follows:

Rule 1: If the phase B slope is less than or equal to one word gain per week, draw the goal line at a slope of 1.0 throughout the remainder of phase C.

Rule 2: If the phase B slope is greater than 1 word gain per week, the goal line is drawn as a continuation of the phase B slope.

It should be noted that the criterion slope of 1.0 words per week was chosen as a minimum goal since this was the mean baseline slope during year 1 of the study.

Once the goal line was correctly drawn for each of the "goal oriented" teachers, they were given the criteria for referencing progress to the goal and making a change where necessary. It was specified that a change in program was allowed at any time, but also that a program modification must be implemented if four consecutive one minute reading sample measurements fell below the goal line. Thus, two distinct data decision rules were established for comparisons within the C phase of the experiment. Since teachers were randomly assigned to data utilization groups, any differences in phase C mean slopes would indicate the superiority of one decision rule over the other.

As in the analysis for year 1, concern for the equality of variance covariance matrices led to a separate analysis of phase differences within treatments for groups with large enough sample sizes for a reasonable test of the hypothesis of no differences among phases. Another reason for running these separate tests was to determine whether there were overall differences between teachers within a group and whether some teachers presented a different pattern of means across phases for their subjects. The main group of interest for this analysis was the Direct Instruction group which attracted the majority of teachers for initial training and thus provided the greatest number of cases for considering our experimental hypotheses.

A no-between, one-within repeated measures ANOVA on the Direct Instruction group means yielded a probability of .06 that the observed mean differences among phases was due to chance [F (2,90) = 2.51, p = .06]. The mean slopes for phase A and phase B were virtually identical (1.04 and 1.01, respectively), while the mean for phase C was 1.38 (see Appendix H for a complete set of summary statistics).

A test of teacher differences was done with a one-between (i.e. teacher factor), one-within (i.e. phase) repeated measures analysis of variance. The results of this analysis showed that there was a significant overall difference in student performance between teachers [F (14,45) = 2.75, p = .02], but that the interaction between phase and teacher was non-signficant [F (2,90) = 1.43, p = .12]. While the main effect difference between teachers may indicate differential teaching effect, it should be noted that students were not randomly assigned to teachers. Some teachers may simply have worked with higher achieving students, who demonstrated greater slopes of improvement in reading across all phases. The fact that the teacher-phase interaction factor was not found to be significant would lend support to this hypothesis (that is, if baseline conditions were considered to be equivalent across teachers). Failure to obtain teacher by phase interaction may also call into question the long-standing assumption that specific instructional interventions work better for some teachers and not others.

A similar set of analyses was performed on the Precision Teaching group. The mean slopes for each phase were as follows: phase A = 1.35, phase B = .52, and phase C = 1.46. The repeated measures ANOVA found the

differences among these means to be significant [F (2,26) = 3.90, p = .03], with followup tests indicating that the mean for phase B was significantly less than the means for both phases A and C. The test of differences among teachers again showed a main effect for teachers [F (3,10), p = .05] but no significant teacher by phase interaction [F (6,20), = 1.13, p = .38]. Separate tests of the phase and teacher differences within the Cognitive Strategies group were not performed because of limited sample size (i.e. only 4 students).

The student outcome data from year 2 was then reconsidered, putting emphasis on the rate of reading achievement improvement generated during phase C. Regression slopes were recalculated for the entire period up until phase C. This slope, based on a minimum of 12 weeks worth of data, was considered to be a much more stable estimate for comparison with phase C slopes. In addition, we felt that a test of AB vs. C was conceptually warranted as a comparison of the rate of progress, while new interventions, such as precision teaching and direct instruction, were being learned and implemented, the final period, where the "full blown" treatments were in effect with modifications taking place to maximize treatment effectiveness. This analysis also had the advantage of including subjects who had been dropped from the data set because of limited phase A data.

The overall mean slope for the combined AB period with 98 subjects from all five groups included in the analysis was 1.33, while the slope for phase C for this entire group of subjects was 1.53. Analysis of the differences between these two phases, using a dependent t-test, yielded a random sampling probability less than .10 that the differences were

attributable to chance (t = 1.72, p = .08). A similar analysis was performed on the Direct Instruction (DI) and Precision Teaching (PT) groups combined and then separately. The p values associated with these tests were between .10 and about .30. A summary of these results is as follows:

Group	number of subjects	Mean Slopes AB	C	t value	p value
DI & PT	72	1.24	1.46	1.69	.095
DI	57	1.27	1.48	1.39	.171
PT	15	1.13	1.39	1.03	.320

The results above, showing that the mean slope during phase C for these groups was consistently higher than the AB slope, and the prior results, indicating superior performance during phase C in comparison to phase B, are interesting findings in light of the results from year 1. In the year 1 analysis, phase B slopes were substantially higher than baseline slopes and the slopes for phase C interventions. Year 2 results show the opposite trends. The fact that treatment interventions periods were timed differently from year 1 to year 2 (e.g. holiday vacations fell at during different phases) might account for the different pattern of means. However, the superior performance of phase C may be indicative of the different research design during year 2, which allowed teachers to continue the interventions developed during phase B for a greater period of time and to use one of two decision rules in deciding when to make additions or modifications to their interventions on an individual basis.

Data Utilization Outcomes

Subjects within the Direct Instruction (DI) and Precision Teaching (PT) groups were randomly assigned to one of two data utilization conditions

during phase C. The "treatment oriented" group was forced to make some modification or addition to their program plan after each set of 9-12 data points. They determined the degree of change necessary by comparing the slopes of reading progress for successive time periods. The "goal oriented" group referenced their program modification decisions to an aim line drawn at a slope either equivalent to the phase B slope or at a rate of 1.0 words per week, whichever was greatest. If four consecutive measurements fell below this goal line, a change in program was mandated.

A completed crossed two-way ANOVA was performed to determine if the population means associated with performance during phase C for the two data utilization strategies could be considered to be different. The means for each treatment group and data utilization strategy, along with the complete ANOVA table, are presented in Table 12. The results indicate a main effect difference in favor of the "treatment oriented" aproach which would be obtained in repeated random samples of the same size less about 6% of the time by chance ($F = 3.57$, $p = .06$). The group main effect and treatment by group interaction were found to be non-significant.

Insert Table 12 about here

We feel that the results lend support to the "treatment oriented" approach which utilizes time series slopes of progress for deciding on program modifications. With replication of these findings, an argument could be made for employing this strategy when using continuous curriculum-based measures.

Data on the number and type of changes made by each treatment group indicated that the teachers in the treatment oriented group made more frequent changes than the goal oriented group by a ratio of about 3 to 2 (i.e. 149 changes by treatment oriented group and 102 by the goal oriented group). Teachers also evaluated the effectiveness of these changes using the criteria appropriate to each strategy. That is, the goal oriented group considered an intervention effective if the data remained above the goal line while the treatment oriented group made decisions based on quarter-intersect generated slopes. Interestingly, the mean percentage of effective interventions was nearly identical for both groups, with goal oriented interventions effective 59% of the time and treatment oriented interventions effective 68% of the time.

One can speculate that the overall superiority demonstrated by the treatment oriented strategy may be a function of the greatest number of interventions, a more stringent criterion for success, or a combination of these factors. Either of these factors might be limitations of the more standard goal oriented approach to evaluating individual education programs (IEPs). The goal oriented strategy closely resembles the standard practice in special education of setting a long term goal at the beginning of the year which serves as the chief reference point for assessing student achievement and instructional success. As long as the student is achieving at or above their goal level no changes in program are considered. The strategy of referencing the gains during current instructional periods with previous periods, in order to determine whether individualized interventions could produce superior performance, may produce achievement

which is consistently higher than the goal expectations. This hypothesis should be tested in future research.

Teacher Directed vs. Computer Directed Outcomes

The Edmark Reading Program and ESWP/Computer Software interventions were designed to test whether these basic whole word curricula would produce better reading achievement gains than baseline and additionally whether the computer added significantly to the progress made over and above a teacher directed approach to instruction.

Separate tests of the Edmark and ESWP interventions were performed in order to see whether the mean slopes for phases A, B, and C differed from each other. With a sample size of seven students, the random sampling probability associated with differences among the phase means was greater than .25 in a repeated measures ANOVA. Regression slopes were also calculated for the entire B-C period for the Edmark intervention. The mean of the BC slopes (2.195) was substantially greater than the mean slope for baseline (1.172). However, because of such a small sample size and inadequate power to find real differences, the paired t-test with 6 degrees of freedom, the p value for this test was again greater than .25.

Similar tests were performed to estimate the probability of differences among the three phases for the ESWP/Computer software intervention. The repeated measures ANOVA yielded an F ratio of 1.00 and p value greater than .25. The matched pairs t-test, with 12 degrees of freedom, between phase A and the combined BC phase yield a t value of .45 and p value of .66.

Changes between phases for the Edmark and ESWP interventions were counterbalanced with the change from teacher directed to computer directed

instruction within each group. This was the essential contrast to be tested statistically. The mean slopes for the computer directed instruction were 1.68 for the Edmark Reading Program packaged computer program and 1.55 for the ESWP with individualized computer software lessons. The mean slopes of improvement for the teacher directed interventions were somewhat higher for the Edmark group (2.01) but somewhat lower for the ESWP group without computer software (mean slope = 1.45). The repeated measures ANOVA comparing these means was nonsignificant for the main treatment effect difference between teacher directed and computer directed ($F = .09$, $p = .77$) and for the group by treatment interaction ($F = .522$, $p = .49$). These results are consistent with teacher comments that the slopes of improvement appeared to be no different between the two methods. They also noted that the gains in student reading performance did not commensurate with the amount of time required to develop individualized software programs for the ESWP interventions.

Teacher Knowledge and Attitude Outcomes

Data from the control teachers for year 2 had not been coded and entered on computer at the time of this report. The preliminary results presented here are based solely on the data from year 2 experimental teachers. The data from the Intervention Knowledge instrument and the Belief Scale were analyzed with a paired t-test to detect the occurrence of significant differences between the pretest and posttest measures.

In accordance with year 1 results, the experimental teachers for year 2 significantly increased ($t = 8.20$, $p < .001$) the number of interventions listed in response to short case study examples of students with reading

problems. The pretest mean was 15.88 and the posttest mean was 19.96. On the rating of effectiveness of these interventions these teachers showed an increase ($t = 1.95$, $p = .06$) from a mean rating of 3.63 to a mean rating of 2.96 on the five point Likert Scale. This change was, interestingly, in the opposite direction from year 1 results.

Results of the Belief Scale indicated that mean estimates of the percentage of students positively affected by the listed interventions and mean ratings of judgment certainty changed little from pretest to posttest measurement. The mean percentage estimate increased slightly from 63.6% to 64.6%. The mean certainty rating also increased marginally ($t = .63$, $p > .30$) from a mean of 2.31 to a mean of 2.19, on the scale where a rating of 1 represented greatest certainty and 5 least certainty.

Without a complete analysis based on comparison with the control group, a thorough explanation of these findings is not possible. Speculation on the possible reasons for the different pattern of means obtained in the year 1 analysis centers on the different design of training and student outcome evaluation employed for year 2.

Year 1 teachers were restricted to the evaluation of specific intervention contrasts in a fairly highly controlled study where individual innovations from the standardized procedures was highly discouraged. Evaluation of these specific contrasts were aggregated and presented so that all teachers became aware of the results which may have been contrary to their personal predictions.

On the other hand, year 2 teachers were encouraged to innovate in finding individualized interventions which would modify or supplement the

standard procedures of their particular treatment. Given the freedom to make such innovative changes, the number of different types of interventions in which they could have direct experience increased. The fact that teachers were able to experiment more freely to find effective alternative treatments increased the probability that they would find some intervention variables which were effective for their individual students.

The increase in ratings of effectiveness associated with the teachers' own list of intervention alternatives on the Intervention Knowledge Scale may have been a function of the generally more positive outcomes evidenced by their students. The fact that phase C interventions were in general less successful for year 1 students and generally much more effective for year 2 students may be an important factor in the overall ratings of effectiveness.

A thorough content analysis of the lists of interventions from the Intervention Knowledge instrument has not been done. It would be interesting to note whether year 2 teachers were more likely to list intervention modifications which were made during phase C for their own students. Some of the minor modifications involving motivational variables (e.g. changes in reinforcers, point values, short term contingencies, etc.) may have been more frequently listed for year 2 teachers. Year 1 teachers may have been more likely to list overall program interventions such as peer tutoring, precision teaching, direct instruction etc.

Another possible explanation for the differences concerns errors of measurement associated with continued exposure to the same instruments. The fact that most year 2 teachers were administered the same instruments 5

times may have affected the sensitivity of the measure. One could suggest that preliminary exposure to the instrument before any experience with training relevant to the content of the instrument, might affect the expectations associated with subjective ratings. If, however, the measures are considered to be valid indicators, one would have to conclude that the year 2 teachers did not demonstrate the attitude changes expected by the investigators.

Summary of Outcomes

Analysis of two years of research data indicates clearly that Special Education Resource Teachers were successfully trained to identify more effective programs for their students. Teachers, who were instructed in the components of curriculum-based measurement, developed skills in data based assessment of student performance which allowed them to quantitatively compare the rate of student reading achievement under contrasting instructional interventions. Systematic training in treatment alternatives allowed teachers essentially to conduct single subject experiments comparing performance during baseline conditions with experimental phases. Training in the quarter-intersect method of estimating slopes of progress on graphed data enabled teachers to make direct comparisons of successive experimental phases. Teachers evaluated changes in the rate of student achievement as a function of instructional interventions for individual students and then aggregated results to get a generalized assessment of treatment efficacy.

The knowledge of results from continuous evaluation of student progress not only provided feedback on the relative effectiveness of a common set of

interventions, but also allowed teachers to evaluate program modification from an idiogaphic perspective (i.e. answering the question of how best to tailor programs to meet individual student needs). Results from the measurement of teacher knowledge and beliefs indicates that training in experimental teaching can effect teacher beliefs related to the effectiveness of educational interventions. Clearly, year 1 teachers decreased their estimates of the number of students who would be positively effected of standard educational treatments and decreased subjective ratings of the effectiveness of interventions proposed by themselves to deal with specific reading problems as a function of participation in the project. Yet, the different pattern of results from the first to second year of the project suggests that the specific type of training and direct experience provided during year 1 (i.e. differing from year 2) may be determinitive of this effect.

A change in the design of the study during year 2 allowed the investigators to evaluate the relative effectiveness of two differnt approaches to the utilization of continuous curriculum-based measures. The results indicated greater increases in student performance under a "treatment oriented" strategy where frequent program modifications were made on the basis of a direct comparison of progress slopes in successive intervention phases. The apparent superiority of this set of decision rules over the more traditional "goal oriented" strategy has important practical implications for policy decisions regarding the establishment of academic goals and the criterion for assessing goal attainment.

Finally, this collaborative research and inservice training project between the Minneapolis Public Schools and the University of Minnesota has

fostered the professional growth of a large number of special education personnel. The Special Education Resource Teachers, teacher trainers, special education administrators, and project staff who participated in the Experimental Teaching Project have developed knowledge and skills which will enable them to be more effective in meeting the needs of special education students.

Table 1

Intervention Knowledge: Number

	Group 1 (Exp)	2 (Cont 1)	3 (Cont 2)	Marginal Means	F	Probability
1	17.00	17.62	13.50	16.02	Group 8.59	.00 *
2	19.09	18.00	15.55	17.57	Time 20.31	.000 *
3	22.77	20.61	16.25	19.95	Group x Time 1.50	.21
Marginal Means	19.62	18.60	15.10	17.78		

Table 2

Intervention Knowledge: Effectiveness

	Group 1 (Exp)	2 (Cont 1)	3 (Cont 2)	Marginal Means	F	Probability
1	3.70	3.68	4.07	3.82	Group 3.82	.08
2	3.84	3.84	3.80	3.83	Time .08	.93
3	3.48	3.74	4.15	3.79	Group x Time 2.64	.04 *
Marginal Means	3.67	3.75	4.01	3.81		

Scale: 1 = not effective

5 = highly effective

Table 3

Belief Scale: Mean Percent

	Group 1 (Exp)	2 (Cont 1)	3 (Cont 2)	Marginal Means	F	Probability
1	62.45	62.04	64.19	62.92	Group 1.85	.16
2	67.95	67.56	64.86	66.33	Time 15.30	.000 *
3	49.82	61.69	68.13	58.94	Group x Time 15.91	.000 *
Marginal Means	60.08	63.76	65.73	62.90		

Table 4

Belief Scale: Certainty

	Group 1 (Exp)	2 (Cont 1)	3 (Cont 2)	Marginal Means	F	Probability
1	2.211	2.04	2.20	2.16	Group 1.99	.15
2	2.52	2.35	2.43	2.45	Time 5.56	.005 *
3	2.63	2.15	2.29	2.40	Group x Time .96	.43
Marginal Means	2.45	2.18	2.31	2.33		

Table 5

Mean Slopes for Phases A, B, and C by Contrast Groups

(Number of data points per phase > 11)

Group	Baseline	Phase B	Phase C	Subjects
1	1.007	(Holt, Whole Word Voc.) 1.703	(Holt Phonics Vocab.) .963	N = 12
2	.559	(Holt Basal Whold Wd.) 2.171	(Merrill Linguistic) .749	N = 16
3	.239	(Language Experience) .019	(Tokens & Oral Reading) .724	N = 13
4	.961	(Hi-interest/Low Vocab.) 1.248	(Language Experience) .173	N = 8
5	1.419	(C.Age Tutor/Phonics) 1.922	(C.Age Tutor/Free Read.) 1.311	N = 10
6	.473	(Indiv. Reinforcement) 1.816	(Group Reinforcement) 1.573	N = 9
7	.935	(Prec. Teach w/o Rate Bdg.) 1.341	(Prec. Teach & Tokens) 1.992	N = 12
8	1.171	(Free Choice Reading) 1.104	(Language Experience) .880	N = 15
TOTAL	.832	1.407	1.039	TOTAL N = 95

Table 6

Repeated Measures Analysis of Differences
Among Phases A, B, and C Within Groups

S.V.	Df	Sums of Squares	Mean Square	F	Probability
Bet. Groups	7	37.1391	5.306	2.77	.012
Error 1	87	158.7211	1.824		
Bet. Phases					
w/Group 1	2	4.1337	2.067	1.36	.259
w/Group 2	2	24.8245	12.412	8.17	.0004
w/Group 3	2	3.3845	1.692	1.11	.331
w/Group 4	2	4.9642	2.482	1.63	.198
w/Group 5	2	2.1315	1.066	.70	.498
w/Group 6	2	9.2100	4.605	3.03	.050
w/Group 7	2	6.8193	3.410	2.24	.109
w/Group 8	2	.6951	.348	.23	.796
Error 2	174	264.5061	1.520		
TOTAL	281				

Followup Tests Within Group 2

Baseline vs. Phase B (Holt Basal Whole Word)

t = 3.03 p = .001

Baseline vs. Phase C (Merrill Linguistic)

t = .37 p = .358

Phase B vs. Phase C (Holt Basal Whole Word vs. Merrill Linguistic)

t = 2.72 p = .004

Followup Tests Within Group 6

Baseline vs. Phase B (Individual Reinforcement)

t = 2.57 p = .006

Baseline vs. Phase C (Group Reinforcemnt)

t = 2.11 p = .018

Phase B vs. Phase C (Individual vs. Group Reinforcement)

t = .46 p = .250

Table 7

Within Phase Comparisons Among Treatment Groups

S.V.	DF	Sums of Squares	Mean Square	F	Probability
Bet. Groups					
W/Phase 1	7	12.7896	1.826	1.12	.357
W/Phase 2	7	61.2283	8.869	5.43	.002
W/Phase 3	7	23.2852	3.326	2.05	.057
Pooled Error	56		1.622		

Followup Tests Within Phase 2
Comparison of all Possible Pairs of Means
(Only Significant Differences Listed)

Holt Basal Whole Word (Group 2) vs. Language Experience (Group 3)

t = 3.39*

Cross-Age Tutor with Phonics (Group 5) vs. Language Experience (Group 3)

t = 3.59*

Individual Reinforcement (Group 6) vs. Language Experience (Group 3)

t = 3.39*

Holt Whole Word Vocabulary (Group 1) vs. Language Experience (Group 3)

t = 3.12*

*Indicates a significant difference using the Bonferroni critical value for 28 pair-wise comparisons with 56 degrees of freedom, at the .05 level of significance = 3.01.

Table 8

Phase B and C Treatment Comparisons

(Percentage of Students Better Under B, C, and No Difference, Using a 1.25 slope Difference Criterion)

	Phase B		Phase C
Group 1	Basic Whole Word 61.7	No Diff. 61.7	Basic Phonics 66.7
Group 2	Basic Whole Word 52.9	55.9	Nonstd Linguistic 11.8
Group 3	Language Experience 55.1	30.8	Oral Reading & Tokens 27.3
Group 4	Mi-Skill Ln Wsp. Relsh 66.6	30.9	Language Experience 22.2
Group 5	Tutor w/Phonics 88.3	66.7	Tutor w/Free Reading 0.0
Group 6	Indiv. Reinforcement 27.3	63.3	Group Reinforcement 27.3
Group 7	Prac. Teach w/o Rule 8.3	30.0	Prac. Teach w/Tokens 63.7
Group 8	Free Choice Reading 58.8	60.0	Language Experience 12.5

Table 9

Effect of Familiarity with Student and
Use of a Judgmental Aid on Inter-Observer
Agreement of Decisions Based on the Graphed Data

Source of Variation	Sum of Squares	df	F	p
Between Subjects				
Certainty (C)	.188	1	.36	---
Within Cells	1.000	22		
Within Subjects				
Judgmental Aid (JA)	.321	1	10.39	< .01
C x JA	.302	1	9.81	< .01
Within Cells	.030	22		

Table 10

Criteria Used in Judging
the Effectiveness of Treatment

Type of Analysis	Level	Number of Graphs Using Each Criterion: Slope	Variability	Combination
Pre-Training				
Teacher	15	37	20	31
Reliability Observer	17	41	23	21
Post-Training				
Teacher	3	70	5	18
Reliability Observer	6	80	4	13

TABLE 11 - YEAR TWO

MEAN SLOPES FOR PHASES A,B AND C BY CONTRAST GROUPS

(Number of data points per phase >11)

	Group	Baseline	Phase B	Phase C	Subjects
1	Direct Instrution	1.04	1.01	1.38	N= 46
2	Precision Teaching	1.35	.52	1.46	N= 14
3	Cognitive Strategies	1.85	1.28	.93	N= 4
4	Edmark Reading Prog.	1.17	1.69	2.10	N= 7
5	ESMP/Cosputer	1.46	1.23	1.78	N= 12
	TOTAL	1.20	1.03	1.49	N= 83

Repeated Measures Analysis of Differences
Among Phases A,B and C by Group

S.V.	Df	Sums of Squares	Mean Square	F
Between Subjects				
Bet. Groups	4	7.845	1.961	1.393
Error 1	78	109.805	1.408	
Within Subjects				
Phase	2	9.083	4.541	5.274 **
Group x Phase	8	8.857	1.107	1.285
Error 2	156	134.383	.861	
TOTAL	248			

** Indicates significance at the .01 level.

TABLE 12

MEAN SLOPES and ANOVA for
DATA UTILIZATION STRATEGIES BY GROUP

Group	Goal Oriented	Treatment Oriented
Direct Instruction	1.27 (29)	1.67 (33)
Precison Teaching	1.19 (8)	1.37 (11)

(Number of students in each cell are in parentheses)

TWO WAY ANALYSIS OF VARIANCE

S.V.	Df	Sums of Squares	Mean Square	F	p
Bet. Groups	1	.640	.640	.935	.337
Strategies	1	2.445	2.445	3.573	.062
Group x Strat	1	.165	.165	.242	.624
Residual	77	52.690	.684		
TOTAL	80				

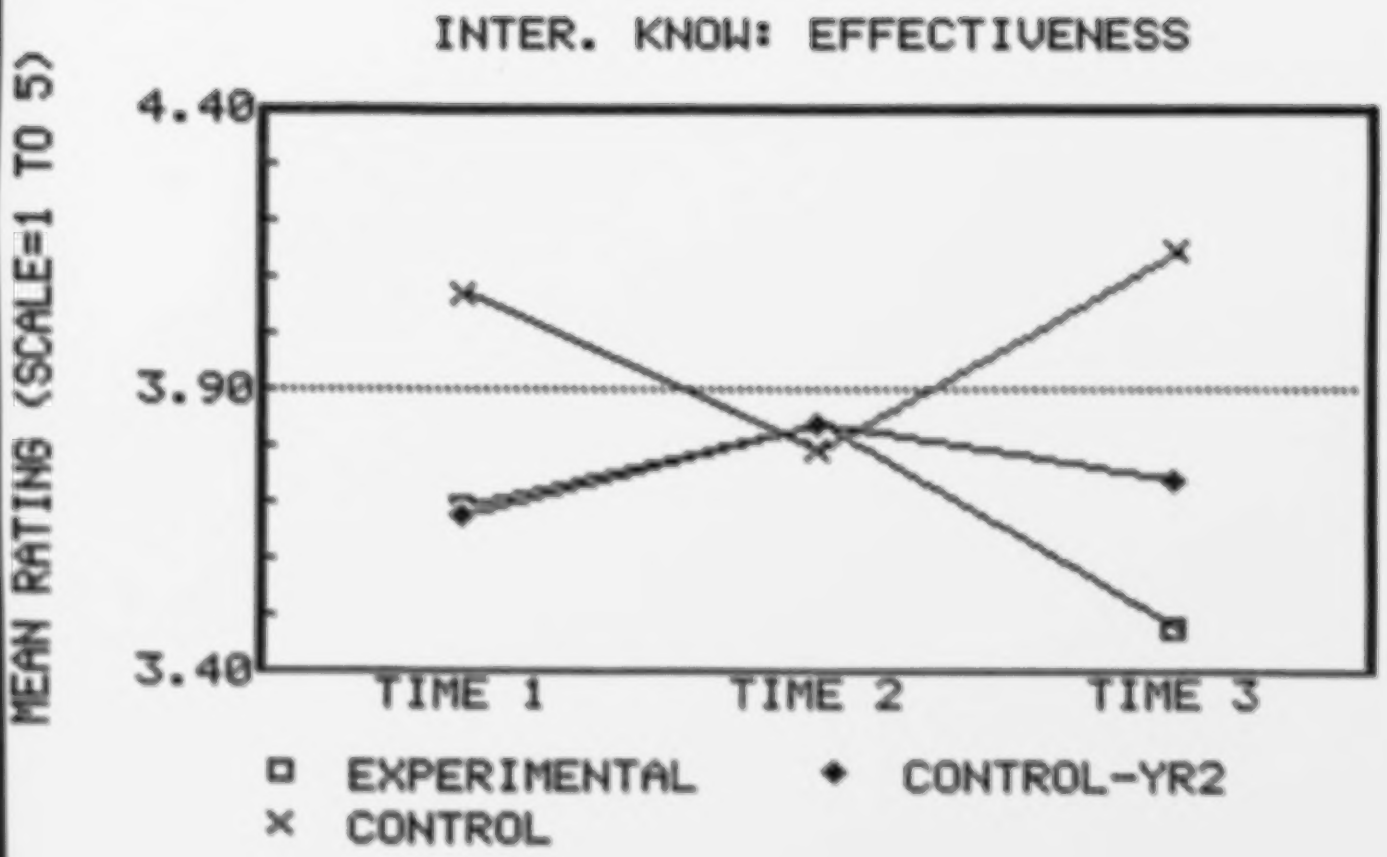
INTER. KNOW: EFFECTIVENESS
MEAN RATING (SCALE=1 TO 5)
4.40
3.90
3.40
TIME 1
TIME 2
TIME 3
EXPERIMENTAL
CONTROL-YR2
CONTROL

Figure 1.

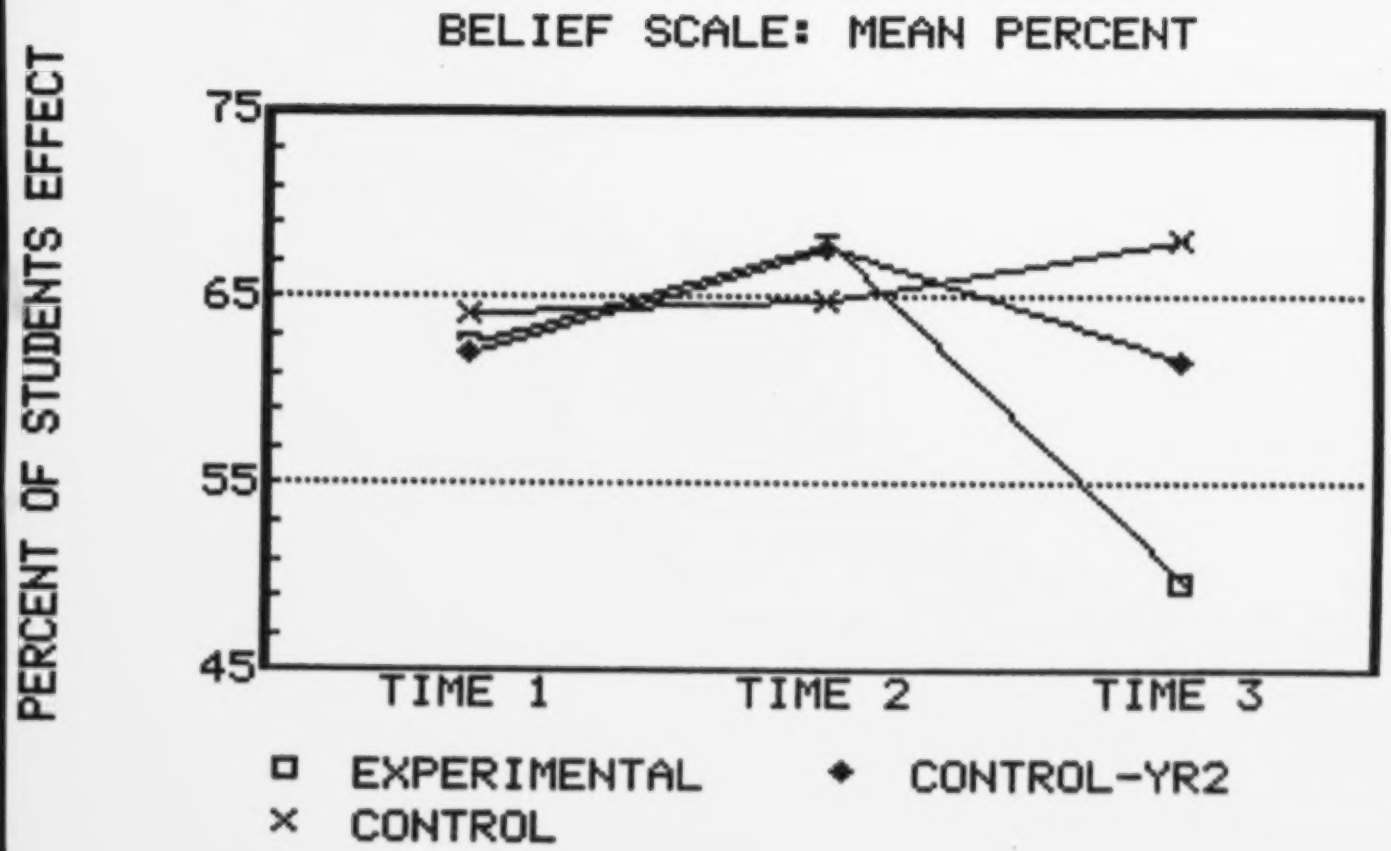
BELIEF SCALE: MEAN PERCENT
PERCENT OF STUDENTS EFFECT
75
65
55
45
TIME 1
TIME 2
TIME 3
EXPERIMENTAL
CONTROL-YR2
CONTROL

Figure 2.

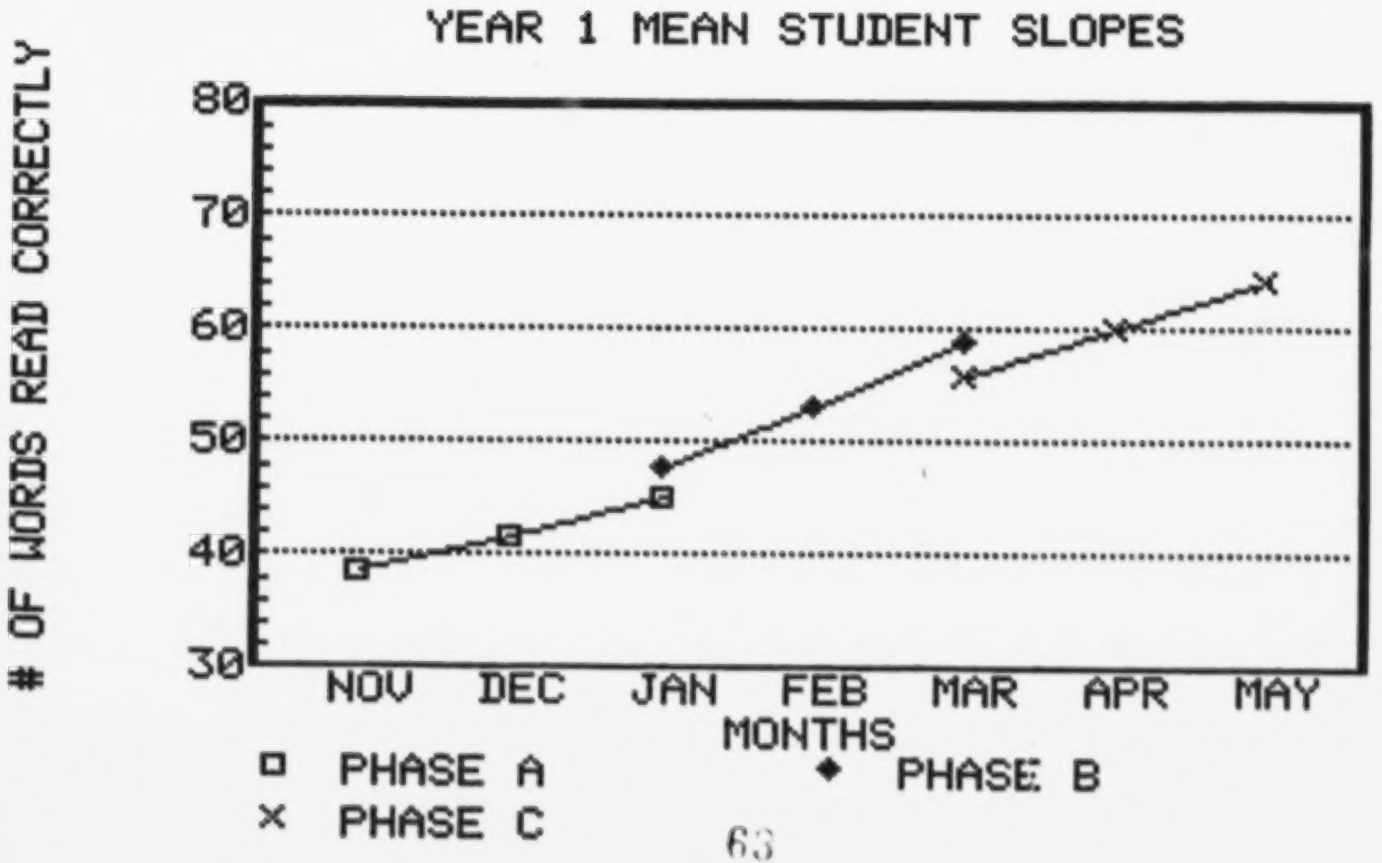
YEAR 1 MEAN STUDENT SLOPES
OF WORDS READ CORRECTLY
80
70
60
50
40
30
NOV
DEC
JAN
FEB
MAR
APR
MAY
MONTHS
PHASE A
PHASE B
PHASE C

Figure 3.

	PHASE B	PHASE C		PHASE B	PHASE C
GROUP I	HOLT WHOLE WORD	HOLT PHONICS	GROUP 5	CROSS-AGE TUTOR (PHONICS)	CROSS-AGE TUTOR (FREE READING)
GROUP 2	HOLT WHOLE WORD	MERRILL LINGUISTIC	GROUP 6	INDIVIDUAL REINF.	GROUP REINF.
GROUP 3	LANGUAGE EXPERIENCE	TOKEN SYSTEM & ORAL READING	GROUP 7	PRECISION TEACHING W/O RATE BUILDING	PRECISION TEACHING WITH TOKEN SYSTEM
GROUP 4	HIGH INTEREST, LOW VOCAB (DOLCH WORDS)	LANGUAGE EXPERIENCE	GROUP 8	FREE CHOICE READING ACTIVITIES	LANGUAGE EXPERIENCE

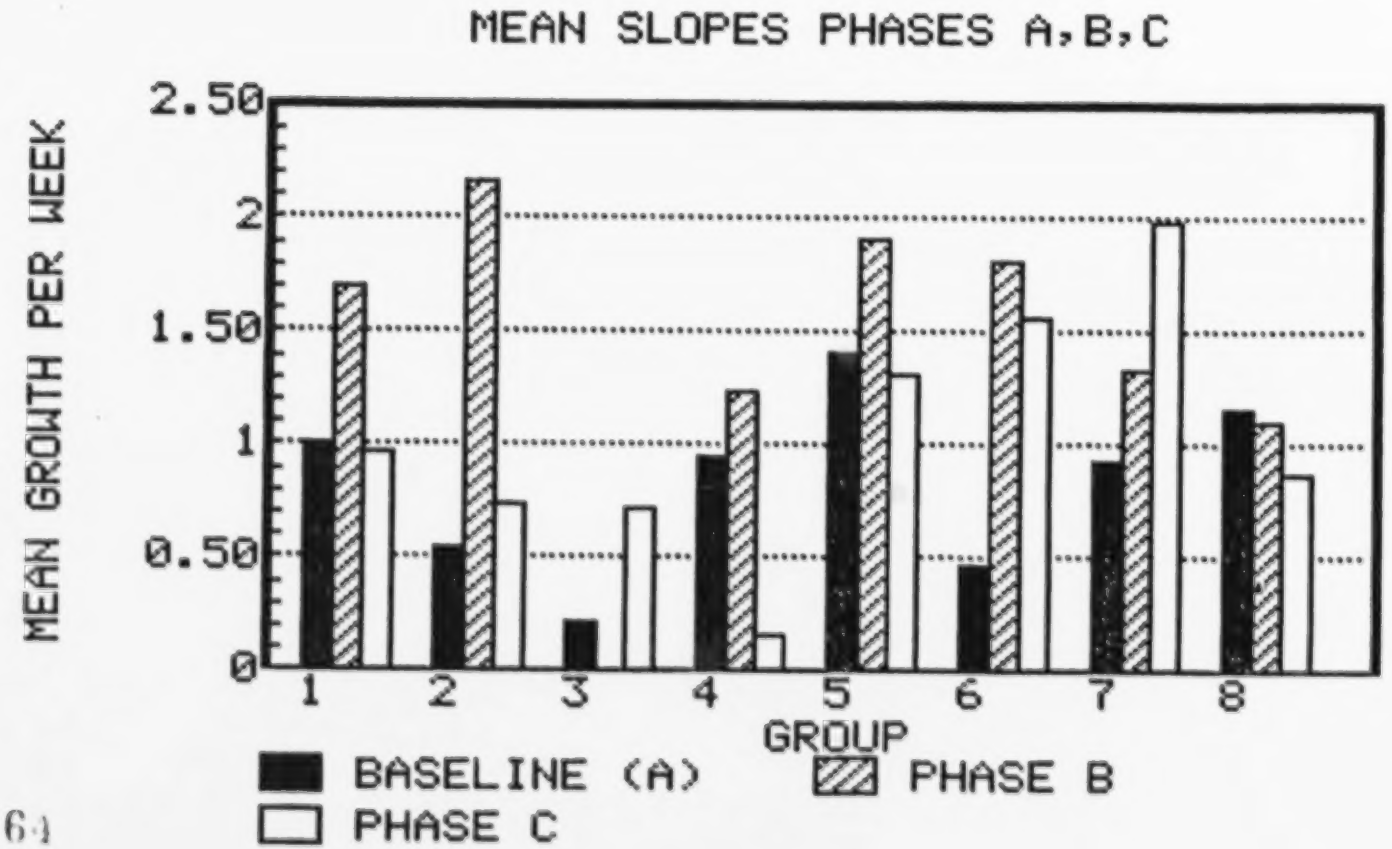

Figure 4.

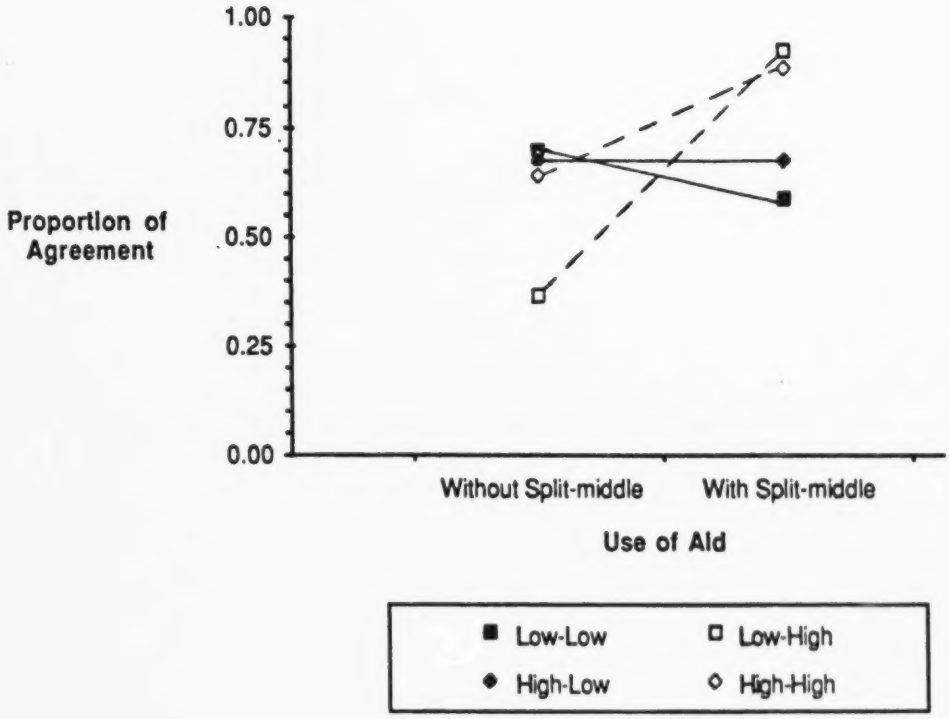

Figure 5. Proportion of graphs agreed on by teacher and reliability observer. Low-Low = Certainty of judgment low in both phases; High-Low = Certainty of judgment decreased with introduction of judgmental aid; Low-High = Certainty of judgment increased with introduction of judgmental aid; High-High = Certainty of judgment high in both phases.

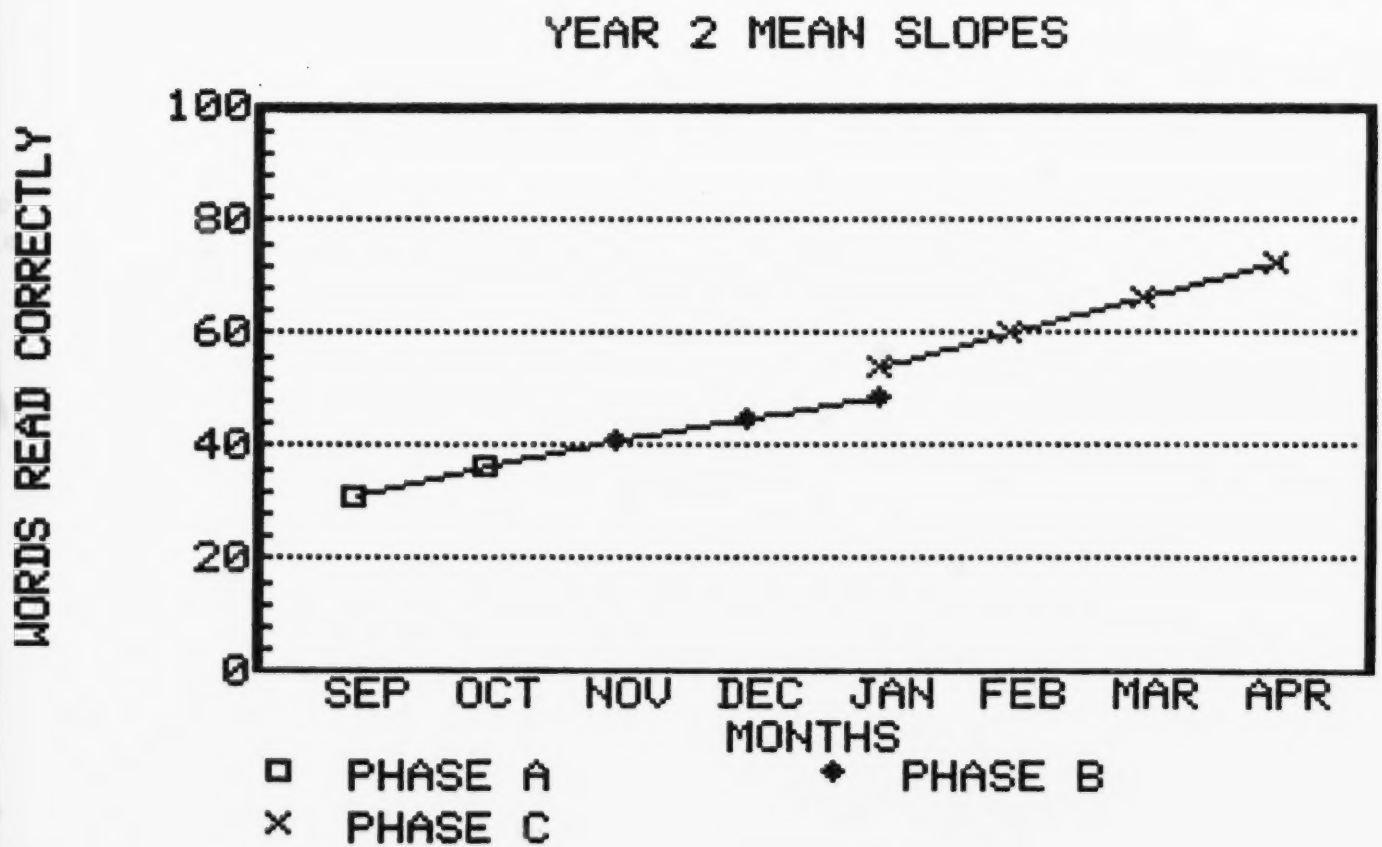
YEAR 2 MEAN SLOPES
WORDS READ CORRECTLY
100
80
60
40
20
0
SEP
OCT
NOV
DEC
JAN
FEB
MAR
APR
MONTHS
PHASE A
PHASE B
PHASE C

Figure 6.

Appendix A

Listing of Instructional Contrasts Generated by SERTs

EXPERIMENTAL TEACHING PROJECT-TEACHER CONTRACTS

1.	Rate criterion vs. accuracy criterion (peer tutoring)	0
2.	PA Gx vs VI Gx (tape)	0
3.	Basal reading vs. direct instruction	0
4.	Daily Gx vs. weekly Gx	0
5.	Language exp. vs free choice reading	0
6.	Reading aloud to adult (not teacher vs Gx) vs. reading silent vs Gx	0
7.	Change to higher level vs. lower level	0
8.	Change to higher level vs. student choice	0
9.	One story contrast vs. tangible Gx	0
10.	Testing vs sight vocabulary vs. accuracy	0
11.	Academic vs comp. vs Gx vs. non-academic Gx (computer)	0
12.	Corrected SS to phonics vs. teacher made instruction	0
13.	Teacher Gx vs. self Gx	0
14.	Daily adult interaction vs. daily alone	0
15.	Daily sound drills vs. no phonic review	0
16.	Testing on daily vocab. vs. no testing	0
17.	Accuracy criterion vs. rate criterion (basal)	0
18.	Gx mastery to instructional level vs. Gx grow to IRI goal	0
19.	Token Gx (tactile) vs. language exp. (contractual)	0
20.	Oral reading 3 min. daily vs. oral reading 10 min.	0
21.	Oral reading alone vs. writing alone	0
22.	Testing in every story vs. testing in every other	0
23.	Testing daily vs. alone	0
24.	Gx every contract vs. intermittent	0
25.	Language exp. vs linguistic	0
26.	Comparative vs. individual token econ.	0
27.	Daily vocab. grades vs. tokens	0

28.	Helps to curriculum based on 1 day vs. 2 day sessions	8
29.	Necessary DV age vs. DSD age	9
30.	Direct teaching vs. individual work (workbook)	4
31.	Oral reading vs. silent	23
32.	Being taken vs. being a tutor	3
33.	Which words by teacher vs. which by peer	5
34.	Direct phonics instruction vs. reg. struct. phonics	2
35.	Social DV to peer tutor vs. tangible RV to PT	2
36.	Consult. + level III vs. level III alone	3
37.	Some were removed fm. activity RV	2
38.	Knowledge impression vs. cognitive struct'n	4
39.	Teacher direction small group vs. peer and/or cross age tutor	3
40.	Teacher led vocabulary review vs. language master (Inst.)	4
41.	3 min. oral reading vs. 6 min.	4
42.	Computer drill vs. teacher drill	4
43.	Teacher chart vs. student chart (Incr. successful)	24
44.	S/S vs. oral reading	1
45.	Oral reading vs. reg. struct. oral reading	1
46.	Reg. struct. level I vs. level I	1
47.	Oral reading alone vs. oral reading with tape	0
48.	Social RV vs. token RV	6
49.	RV for academic only vs. RV for academic and behavior	1
50.	Rule trade reg. & spec. eff. 1 hr./day 1 group vs. own group	0
51.	Reading growth toward DIP goal vs. Reading task completion	0
52.	Level 3 vs. Level 2	1
53.	Blend vs. long sep. (Remedial)	3
54.	Cross age vs. same age tutor	3
55.	Difference 3 oral reading & self here vs. 3 min of engaged reading (no reading)	2

55.	Phonics vs. sight vocabulary	0
56.	Sf. word vs. long. exp.	1
57.	Unit vocabulary with pass vs. Unit vocabulary with language master	1
58.	Computer as Sr vs. free choice as Sr	1
59.	Whole day Sr vs. resource room Sr	1
60.	Student goal setting vs. teacher goal	4
61.	Student chooses curriculum vs. teacher chooses curriculum	0
62.	Precision teaching vs. Direct	0
63.	Silent reading vs. choral reading	1
64.	Choral vs. oral	3
65.	Finding Unit vs. finding Wheeler (Patty)	1
66.	Accuracy Sr vs. complete Sr	0
67.	A.M. peer tutor vs. P.M. peer-time peer tutor	0
68.	Student verbalizes day before vs. teacher review	2
69.	10 min. rehearsal Unit vocab. vs. 2-3 min. sessions (Unit)	1
70.	Linguistic vs. precision teaching	0
71.	Unit vs. long. exp.	7
72.	Quick cycle phonics vs. long term phonics	1
73.	Level II radio. vs. Level II vs class (including students)	0

Appendix D

Description of Instructional Contrasts for 8 Groups

Contrast: FREE CHOICE READING ACTIVITIES VS. LANGUAGE EXPERIENCE APPROACH

Investigators: Nancy McKinley - Lyndale
Ann Johnson - Burroughs
Chris Miller - Andersen Contemporary
Palmer Bridges - Andersen Contemporary

B Phase:

Materials: library books, tapes with read along books, reading games, student reading materials brought from home, comic books, catalogs, menus, tape recorder

Size and Constitution of Group: same as baseline

Time: 15 minutes/day; 5 days/week

Procedure: Each student chooses an activity from teacher-made menu:
1. Read library books or other reading materials.
2. Read into tape recorder.
3. Read with a partner.
4. Listen to story with read-along book.
5. Reading game which includes oral reading of words.

Student charts his daily choice. Teacher sets ground rules: 2 minutes to make choice, no changing choice. Teacher monitors activities and insures on-task behavior.

C Phase:

Materials: large chart paper, markers, word cards, student-made blank books to contain stories, physical objects for motivation

Size and Constitution of Group: same as baseline

Time: 100 minutes/week

Procedure: Day 1
1. Talking about concrete object (5 senses words, classification) teacher writes words as verbal vocabulary (lists)
2. Write story (teacher writes) (on wall) beginning-middle-end or one kid says each sentence. Get kids' names in-motivate.
3. Read story as group.

Day 2
1. Reintroduce concept/vocabulary.
2. Read story - expression, whole group
3. Individuals read story or sentence
4. Isolate on words to study (use window frame) and write on separate list, find word back in story, structural analysis (all words w/s at beginning, etc.)

Day 3 and Day 4 — Drill on individual words (sounding out, matching, have kids make w.s., etc.), they make word cards for games, etc. Phonics and structural analysis.

Day 5 — Re-read story, creative writing, cut-up story, play games, read library books, etc. Copy story into own book.

Week 2 - Same but kids need 5 minutes practice on old words. Time on old stories or words.

Contrast: MOTIVATIONAL STRATEGIES WITH ORAN READING VS. LANGUAGE EXPERIENCE APPROA

Investigators: Cleo Oleson - Jefferson
Cathy Davis - Kenwood
Bonnie Rowell - Kenwood

Phase B:

Token Economy: Points will be awarded for the following behaviors:

1. on time
2. task completion
3. attitude
4. increase in words/minimum goal material
5. following along during oral reading - on task
6. neatness
7. accuracy

Oral Reading: Students will read aloud (or follow along) from instructional material for 20 minutes/day.

Phase C:

Language Experience Approach to reading (5 day procedure as described by Derrill.

Contrast: SIGHT APPROACH VS. PHONICS APPROACH

Investigators: Renaye Picard
Karen Brown
Linda Balko
Mary Shuster

B Phase:

Materials:
1. Holt vocabulary lists appropriate to instructional level
2. word cards
3. games
4. stopwatch
5. worksheets
6. core list 10 words

Instructional Group: Same

Time Allocated: 20 minutes

Procedure:
1. Pretest to determine new words.
2. Present word.
3. Discriminate between 3 words, 3 times.
4. Say the word.
5. Repeat process once.
6. Practice 4 3-4 word phrases.
7. Games.
8. Timings twice a week using cumulative words (w.p.m.)

C Phase:

Materials:
1. syllable patterns
2. word cards
3. worksheets
4. games
5. sound cards in isolation

Instructional Group: Same as baseline

Time allocated: 15-20 minutes

Procedure:
1. Identify phonetic skills to be taught.
2. Present syllable pattern.
3. Discriminate pattern.
4. See/say activities.
5. Hear/write word/pattern/sound.
6. Decode words.
7. Games - at least 3 times a week.

Contrast: B DOLCH "WHOLE WORD" APPROACH HIGH INTEREST, LOW VOCAL VS LANGUAGE EXPERIENCE APPROACH

Intestigators: Kathy Songle
Annette Smith
Rolf Huss

Phase B:

Materials: Dolch flash cards, Dolch worksheets, high interest low vocabulary books from P.I.C.

Instructional Group: Same as baseline.

Time Allocated: 20 minutes per day, 4 times per week

Procedure:
1. 10 minutes teacher presentation of Dolch words and meanings.
2. 5 minutes - Dolch Worksheet
3. 5 minutes of silent reading in high interest material

Phase C:

Materials: Paper, Marking pen, student generated materials. "See Derrill"

Instructional Group: Same as baseline.

Time Allocated: 20 minutes, 4 times per week.

Procedure:
Day 1 Generate words, write story, list target words and discuss meaning.

Day 2 Read story all ways (per Derrill) Group - Individual and review target vocabulary words.

Day 3 Phonic and structural review of target words.

Day 4 "Culminating Activities: which include - creative writing, library materials, games and student choice - all relating to Day 1 vocabulary.

Contrast: A. 3 ONE MINUTE TIMINGS PER DAY IN MERRILL INSTRUCTIONAL LEVEL WITH CHARTING

B. NO TIMINGS, NO CHARTING IN MERRILL INSTRUCTIONAL LEVEL

C. 3 ONE MINUTE TIMINGS PER DAY IN MERRILL INSTRUCTIONAL LEVEL WITH CHARTING AND MOTIVATIONAL STRATEGIES

Investigators: Cheryl Savage - Holland
Toni Silver - Holland
Anne Hoglund - Bethune

Phase B:

No timings, no charting in Merrill instructional level

Phase C:

A. 3 one minute timings per day in Merrill instructional level with charting

Criterion: K-3: 30 per minute to mastery
4-6: 80 per minute to mastery

Motivation-Token System

In teacher charted timings on Thursdays, tokens will be awarded for w.p.m. Tokens can not be accumulated

K-3	4-6
30 - 1 token	80 - 1 token
35 - 2 tokens	100 - 2 tokens
40 - 3 tokens	125 - 3 tokens
45 - 4 tokens	150 - 4 tokens
50 - 5 tokens	200 - 5 tokens
etc.	225 - 6 tokens
	250 - 7 tokens
	300 - 8 tokens

Menu K-3	Menu 4-6
1 token - sticker	1 token - sticker
2 tokens- Glad-gram	2 tokens- Glad-gram
3 tokens- pencil	3 tokens- pencil
4 tokens- food	4 tokens- food
5 tokens- can/pop	5 tokens- can/pop
6 tokens- toy	6 tokens- toy

1 token for high wpm reader of the week

Contrast: INDIVIDUAL R+ OF ACADEMICS VS. GROUP R+ OF ACADEMICS

Investigators: Harriet Capetz - Wilder Contemporary
Pat Swanson - Wilder Contemporary
Marlys Johnson - Keewaydin

B Phase:

Materials: Same as baseline

Group: Same as baseline

Time: Same as baseline

Academics will be reinforced individually using a token economy (points) for task completion (80%), accuracy of completed work (80%) and reading fluency (individual objectives in instructional materials). Reinforcement menu will be used for exchange of points.

C Phase:

Materials: Same as baseline

Group: Same as baseline

Time: Same as baseline

Academics will be reinforced on a group basis using a token economy (points) for task completion, accuracy of completed work and reading fluency. Reinforcement menu will be used for exchange of points.

Contrast: CROSS-AGE TUTOR (PHONICS) VS. CROSS AGE TUTOR (FREE READING)

Investigators: Cheryl Mulhausen - Webster
Sheryl Evelo - Marcy
Carol Sauro - Wilder

B Phase:

Materials: Same as baseline

Instructional Group: Same as baseline

Time Alloted: Same as baseline

Procedure: Peer tutor (cross-age) comes to resource room for 20 minutes three times a week and works with student on phonetic skills. The first 10 minutes is allocated to practice and timings/charting. Resource teacher will direct the start and stop of the one minute timings. Two timings will be given daily on the same passage in the story last read. The best score is recorded by the tutor. Individual student must reach rate criteria before continuing to the next story.

The second 10 minutes is allocated to written phonics instruction (i.e., spelling sentences, vowel discrimination, writing words in sentences, etc.)

C Phase:

Materials: Reading materials

Instructional Group: Same as baseline

Time Alloted: Same as baseline

Procedure: Student selects book from pool of teacher selected books or student selected reading material. Student and cross-age tutor will sit side by side. Student will read material while the tutor provides words missed.

5 minutes - selection period
15 minutes - free oral reading

Contrast: HOLT BASAL READER APPROACH, USING SIGHT VOCABULARY TECHNIQUES VS. MERRILL CURRICULUM APPROACH USING LINGUISTIC WORD FAMILIES.

Investigators: Velma Pibal - Holland
Deb Johnson - Loring
Teresa Leverett - Holland
Bonnie Zachery - Gordon Center
Karen Flink - Anderson Open

B Phase:

Materials:
1. Holt Vocabulary Sentence Sheets.
2. Holt Reader and T.M.
3. Tape Recorder
4. Sight Word Pack

Size of Group: Same as baseline unless mobility dictates otherwise

Time: 20 minutes daily

Procedure:
1. Drill pack - introduce new words as needed.
2. Teacher directed review using new and previously learned words in isolation.
3. Teacher directs presentations of new words in context.
4. Student reads given selections independently into tape recorder.
5. Teacher/student listen to tape and monitor and correct errors.

C Phase:

Materials:
1. Merrill Readers
2. Word Pattern and Letter Flash Cards
3. Phrase Cards
4. Merrill Workbooks
5. Tape Recorder

Size of Group: Same as baseline

Time: 20 minutes daily

Procedure:
1. Teacher directed intorduction of pattern words.
2. Student pronounces and uses in sentence.
3. Teacher directed practice of phrases cards.
4. Read passages silently and orally or vice versa.
5. Oral read into tape recorder from Merrill passages.
6. Monitor taped passages with student.

Appendix C

Instructionai Plan Documentation Worksheet

Student Name _____

IEP Goal Area _____

Time Available for Instruction _____

Location of Instruction _____

Instructional Plan

Instructional Activities	Arrangement	Time	Materials	Motivational Strategies

Intervention Knowledge Scale

After I finish reading the following case study I want you to list as many intervention strategies as you can think of that would help to increase the student's reading achievement. Please try to limit each answer to one phrase beginning with a verb. For example, for a student experiencing math difficulties you might write, "Give him a number line to use," or "Have a peer assist him on facts using flash cards," or "Let the student use concrete objects such as cuisenaire rods." List any strategy you can think of and don't be concerned about how effective you think it might be. You will have 1 minute to think and then four minutes to write strategies.

Joe has just been placed in a special education resource room due to a severe reading problem. He is a fifth grader in a rural school district. He is reading in the 2nd second grade reader in the BSRR curriculum. He scored at the 15th percentile on a standardized reading measure. He is a word-by-word reader and has much difficulty in answering comprehension questions that go beyond the main idea. In addition, he appears to be losing interest in trying to improve his reading skills noted by the various avoidance behaviors he demonstrates.

Strategies	not effective				highly effective
1.	1	2	3	4	5
2.	1	2	3	4	5
3.	1	2	3	4	5
4.	1	2	3	4	5
5.	1	2	3	4	5
6.	1	2	3	4	5
7.	1	2	3	4	5
8.	1	2	3	4	5
9.	1	2	3	4	5
10.	1	2	3	4	5
11.	1	2	3	4	5
12.	1	2	3	4	5
13.	1	2	3	4	5
14.	1	2	3	4	5
15.	1	2	3	4	5
16.	1	2	3	4	5

(Directions are same as before.)

Susan is in a suburban second grade and is gradually falling farther behind her classmates in the HOLT reading curriculum. She has a great deal of difficulty in sounding out new words. Even when presented with words she previously has been taught (and supposedly mastered), she often seems unable to recognize the words. She has just been evaluated and found eligible for resource service on the basis of an ability-achievement discrepancy.

Strategies	not effective 1 2 3 4 5 highly effective
1.	1 2 3 4 5
2.	1 2 3 4 5
3.	1 2 3 4 5
4.	1 2 3 4 5
5.	1 2 3 4 5
6.	1 2 3 4 5
7.	1 2 3 4 5
8.	1 2 3 4 5
9.	1 2 3 4 5
10.	1 2 3 4 5
11.	1 2 3 4 5
12.	1 2 3 4 5
13.	1 2 3 4 5
14.	1 2 3 4 5
15.	1 2 3 4 5
16.	1 2 3 4 5
17.	1 2 3 4 5

NAME _____ GROUP _____

Directions: Please consider the following teaching strategies and how effective you think they might be over the course of a six week period in producing a change in reading achievement. First, estimate the percentage of resource students with whom you believe the teaching strategy may be effective. Estimate to the nearest decile, so you will be choosing 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% of the students. Second, please rate how sure you are of your answer. Leave blank any strategy with which you are totally unfamiliar.

	% of students it will be affecting	uncertain ... certain
1. teacher models expected behavior	_____	1 2 3 4 5
2. high ratio of feedback to student response	_____	1 2 3 4 5
3. Dister	_____	1 2 3 4 5
4. teaching skills in the order in which the curriculum presents them	_____	1 2 3 4 5
5. one to one instruction (individual tutoring)	_____	1 2 3 4 5
6. cognitive strategy training to improve reading comprehension	_____	1 2 3 4 5
7. peer and/or cross-age tutoring	_____	1 2 3 4 5
8. point systems (token economies) for academic work	_____	1 2 3 4 5
9. immediate feedback and error corrections	_____	1 2 3 4 5
10. praise	_____	1 2 3 4 5
11. self charting of academic performance	_____	1 2 3 4 5
12. teaching strategies based on whole brain theories	_____	1 2 3 4 5
13. teaching strategies based on modality strengths and weaknesses	_____	1 2 3 4 5
14. mastery learning	_____	1 2 3 4 5
15. telling students your behavior objectives prior to instruction	_____	1 2 3 4 5
16. remediation of perceptual problems	_____	1 2 3 4 5
17. practicing aloud or writing rather than silent practice	_____	1 2 3 4 5
18. diagnostic-prescriptive teaching	_____	1 2 3 4 5
19. Precision Teaching	_____	1 2 3 4 5
20. Merrill reading groups	_____	1 2 3 4 5
21. direct service in resource room rather than indirect service provided to the classroom teacher	_____	1 2 3 4 5
22. teacher-lead large group instruction rather than teacher-lead small group instruction	_____	1 2 3 4 5
23. direct instruction approaches	_____	1 2 3 4 5
24. repeated readings to mastery	_____	1 2 3 4 5
25. neurological impress	_____	1 2 3 4 5
26. language experience approaches	_____	1 2 3 4 5

(DWCB)

		% of students it will be affecting	certain				uncertain
27.	vocabulary development	_____	1	2	3	4	5
28.	study skills training	_____	1	2	3	4	5
29.	sound-symbol blending	_____	1	2	3	4	5
30.	involving parents through home-school contracts	_____	1	2	3	4	5

31. Please list any other teaching strategies not mentioned above you believe to be effective, & estimate the proportion of the students.

32. Do you believe that the rate and accuracy of reading words aloud from text is a valid measure of a student's reading achievement? Why, or why not?

Appendix I

Student Outcome Descriptive Data

Group	Phase	Mean	Standard Deviation	N
1	A	1.007	.879	
	B	1.703	1.398	12
	C	.962	1.339	
2	A	.559	1.217	
	B	2.171	1.766	16
	C	.749	.998	
3	A	.239	.899	
	B	.019	1.498	13
	C	.724	1.021	
4	A	.961	1.300	
	B	1.249	1.009	6
	C	.173	1.510	
5	A	1.419	1.099	
	B	1.923	1.559	10
	C	1.311	2.132	
6	A	.473	.959	
	B	1.816	1.162	9
	C	1.573	2.137	
7	A	.935	.716	
	B	1.341	.927	12
	C	1.992	.740	
8	A	1.171	1.344	
	B	1.104	1.130	15
	C	.000	.905	

Appendix G

Experimental Teaching Project

Training Module Sign Up Sheet

As a participant in the Experimental Teaching Project you have up to 4 full release days you may use for Training Modules. After deciding upon which training sessions you wish to attend, please document on this form with an "X".

# of Teachers	Module
17	Direct Instruction (2 days)
8	Peer Tutoring (1/2 day)
5	Edmark (2 days)
5	Computer Software (3 days)
2	SIMS: Task Analytic Approach (2 days)
11	Reading Comprehension & Cognition (1 day)
16	Contingency Management (1/2 day)
14	Precision Teaching (2 days)

_____

Name

Appendix H

Student Outcome Descriptive Data (YEAR 1)

Group	Phase	Mean	Standard Deviation	N
1	A	1.043	1.232	46
	B	1.011	.994	46
	C	1.385	.753	46
	AB	1.273	1.074	57
	C	1.486	.846	57
2	A	1.353	1.250	14
	B	.522	1.105	14
	C	1.463	.883	14
	AB	1.127	.729	15
	C	1.389	.898	15
3	A	1.848	.628	4
	B	1.282	.632	4
	C	.930	.464	4
	AB	1.249	.253	4
4	A	1.172	.938	7
	B	1.687	.892	7
	C	2.102	1.261	7
	BC	2.195	1.047	7
5	A	1.457	.937	12
	B	1.228	.676	12
	C	1.777	1.390	12
	BC	1.334	.399	12

Appendix I

BEST COPY AVAILABLE

EXPERIMENTAL TEACHING PROJECT

CURRICULUM VARIABLE I

Derrill M. Pankow, Group Leader

QUESTION:

If a student is making extremely limited progress in the very beginning level of a basal reading program such as the Holt Basal Readers, can significant gains be made by introducing a carefully sequenced, highly repetitive, sight word program such as the Edmark Reading Program; and if so, does a student demonstrate more significant progress with the Teacher-Directed Edmark Reading Program or with the Computer Software Edmark Reading Program with Echo II Voice Synthesizer, or are both versions equally effective?

MATERIALS NEEDED:

2 computers-Apple II+ or Apple IIe with 48K; single disk drives; color monitors; joysticks.
2 Teacher-Directed Edmark Reading Programs, Level 1
2 Computer Software Edmark Reading Programs, Level 1 (Joysticks and voice synthesizers are included.)

***Available materials are sufficient for 4 Resource Teachers only, unless Resource Teachers are located within the same school and are thus able to share software and computers.

DESCRIPTION OF EDMARK READING PROGRAM:

Edmark Reading Program, Level 1, utilizes a sight word approach to teach a 150 word vocabulary. Vocabulary is practiced in simple phrase/sentence contexts and brief storybooks. Low-level comprehension skills are included. Teachers follow a specific, consistent procedure and script for teaching new words, utilizing the follow-up materials, and correcting student errors. Students are pretested to determine where they should begin this program, then progress at their own rate of learning and retention. Daily record keeping and frequent progress testing are incorporated into the program. Four kinds of lessons/activities are used: word recognition, following directs books, picture/phrase matching, and story reading. Supplementary materials are usually unnecessary.
Of the 150 words taught in Level 1, 74 are found in the Dolch List and101 are found in Holt, Levels 2-8.

APPROPRIATE STUDENTS:

Edmark Reading Program was designed for special education students who have acquired few reading skills and almost no sight vocabulary in more traditional programs. Studen may be of any age, and computer expertize is not necessary for the software edition.

TEACHER TRAINING:

The Edmark Reading Programs are somewhat unique. Teachers will need one full day to become familiar with the materials, necessary teaching strategies, and the software edition before Phase B begins. The full day training session must be followed one week later with a ½ day session in which teachers ddmonstrate sample Edmark lessons and procedures. Demonstrations are seen as necessary to ensure group consistency in teaching procedures, use of provided teaching scripts, student error corrections, and determining when specific lessons must be retaught. (All 4 components are heavily integrated within the Edmark Program. Teachers with Direct Instruction Strategies will find the Edmark methodology easy to follow, use, and understand.) Teachers will use either the Teacher-Directed or Software edition of Edmark for Phase B, then switch programs for Phase C. ½ day retraining will be necessary before Phase C is begun

EXPERIMENTAL TEACHING PROJECT

CURRICULUM VARIABLE II

Derrill M. Pankow, Group Leader

QUESTION:

Do students increase their fluency in Reading when provided with supplemental instruction on groups of words which occur with high frequency in all reading programs and which are theoretically more difficult for students to learn and retain because these words are primarily structural words, those words which are frequently devoid of sensory meaning and evoke no specific mental images when they are considered in isolation; and if so, do students demonstrate more significant progress when this supplementary instruction is provided with a highly structured commercial program with a set format, methodology, and somewhat traditional learning materials or when provided with a wide variety of Teacher-Controlled Software which enables Resource Teachers to utilize unique learning strategies, capitalize upon high student interest, and provide great flexibility in creating individually designed lessons, drill and practice opportunities, etc. with the most current technology?

MATERIALS NEEDED:

Essential Sight Words Programs, Level 1-2 (6 sets available)
First Reading Books-Dolch (2 sets available)
Basic Vocabulary Books-Dolch (2 sets available)
Apple II Computers with color monitors
Softward Programs (2 programs each-minimum available)

1. MicroQuest and Vocab by TIES
2. Master Spell by MECC
3. Word Wizard by MECC
4. Precision Teaching Practice Sheets
5. The Game Show by Advanced Ideas
6. Tic Tac Show by Advanced Ideas
7. Missing Links-English Editor by Sunburst Corp.

***Available materials are sufficient for up to 10 Resource Teachers with some sharing of materials.

DESCRIPTION OF CURRICULUM VARIABLE II:

Time Required: 15 minutes daily added to students current resource time.

Phase A: Students are being instructd in the Holt Basal Reading Series. They are not receiving specialized instruction, drill, or practice in the Dolch Words or any specialized group of high frequency words.

Phase B: Activities in Phase A remain constant, but 15 minutes daily is added during which students work with the teacher solely on the Essential Sight Word Program and extended oral reading in the Dolch Books. Upon taking the Pre-test to determine which Essential Sight Words students need to learn, Resource Teachers will utilize the set procedures for teaching new words, providing review and remediation, administering and interpreting the Mastery Tests, and guiding the oral reading of the stories in the Essential Sight Words Program. Remaining time from each 15 minute period will be used for oral reading from the supplemental Dolch Books.

Phase C: Activities in Phase A will remain constant, but 15 minutes daily is added during which students work with the teacher on the Teacher-Controlled Software activities uniquely designed for the student. Students will take the Pre-test from the Essential Sight Words Program to determine which words the students need to learn. The Resource Teacher will attempt to utilize the computer and software to introduce new words in some form but need not be restricted solely to its use for all introductory activities. The computer software must be used at least 10 minutes daily for the reinforcement, application, drill and practice activities if the contrast is to be valid with Phase B.

APPROPRIATE STUDENTS:

Students must be receiving their primary instruction in a basal reading program such as Holt. If possible, students should have completed at least Level 6 in the Holt Program or the creation of software lessons will be expremely difficult with such a limited vocabulary. Students should have some familiarity with the computer.

TEACHER TRAINING:

Training requirements for using the Essential Sight Words Program require a maximum of 3 hours. Specific procedures for use of this program are carefully described in the Guide. At the end of the training time, teachers will be expected to demonstrate a sample lesson to the other group members to ensure consistency in the program's use.

Training for using the software programs will require 3 full days. Training should not be on consecutive days as teachers will need some additional time to become familiar with the programs to discover their potential and limitations before attempting to design specific lessons/programs. Teachers must be computer literate as after seeing a demonstration of each software program as well as a variety of possible programs which can thus be designed from the software packet, individual lessons need to be created. Sharing of these programs amongst participants will be expected. At the end of the third training day, teachers will demonstrate at least one lesson and have lesson outlines in written form available.

***Denise Griffith has indicated a willingness to be involved in the training with the computer software as she has considerable experience in this area.

138 words found in Holt, Levels 2-

Appendix B Alphabetized Word Lists*

First 100 Words (*Level 1*)

2- a (16)
8- about (63)
8- after (88)
7- all (43)
also (89)
an (37)
2- and (14)
8- any (78)
2- are (9)
8- as (24)
7- at (28)

7- back (95)
7- be (27)
8- been (49)
before (92)
6- but (34)
by (29)

5- can (2)
3- come (4)
could (71)

did (90)
4- do (76)

even (85)

first (77)
3- for (21)
5- from (35)

5- get (3)
4- go (1)

7- had (32)
2- has (50)
7- have (6)
5- he (10)
8- her (42)
7- him (48)
5- his (25)

5- I (30)
7- if (56)
3- in (17)
4- into (64)
3- is (19)
5- it (22)
its (62)

3- like (8)

made (87)
many (91)
7- may (74)
4- me (84)
more (54)
8- most (86)
7- much (97)
must (93)
5- my (79)

7- new (69)
5- no (55)
not (33)
7- now (80)

7- of (13)
4- on (26)
3- only (67)
or (36)
8- other (68)
8- our (82)
4- out (57)
8- over (83)

4- said (59)
3- see (7)
4- she (11)
7- so (58)
7- some (70)
such (81)

than (65)
7- that (18)
2- the (12)
8- their (46)
7- them (66)
7- then (75)
7- there (44)
these (73)
4- they (38)
8- this (31)
through (94)
7- time (72)
2- to (15)

4- up (61)

5- want (5)
5- was (20)
8- way (99)
7- we (47)
well (100)
7- were (41)
5- what (60)
7- when (51)
6- where (96)
which (39)
5- who (52)
5- will (53)
5- with (23)
8- would (45)

6- you (40)
7- your (98)

*rank from Modified Frequency List indicated in parentheses

All 200 Words

a - 2
about - 8
after - 8
again - 5
all - 7
almost
along
also
always
an
and - 2
another - 8
any - 8
are - 2
around
as - 8
at - 7
away - 5

back - 7
be - 1
because - 8
been - 8
before
began - 8
best - 8
better
between
big - 2
both
but - 6
by

came - 5
can - 2
come - 3
could

did - 4
didn't - 5
different
do - 4
does
done
don't - 7
down - 4

each
early
end
enough
even
ever - 8
every - 8

far - 8
felt
few
find - 4
first
for - 3
found - 8
from - 5

get - 5
give
go - 4
good - 4
got - 7
great

had - 7
has - 2
have - 7
he - 5
help - 8
her - 8
here - 4
high
him - 7
himself
his - 5
how - 8

I - 5
if - 7
in - 3
into - 4
is - 3
it - 5
its

just - 8

kind
knew
know - 7

large
last - 8
later
left
let - 7
like - 3
little - 2
long - 8
look - 3

made - 7
make - 4
many
matter
may - 7
me - 4
might
more
most - 8
much - 7
must
my - 5

need
never - 8
new - 7
next - 8
no - 5
not - 4
nothing
now - 7

of - 7
off - 8
old - 7
on - 4
once
only - 8
open - 7
or
other - 8
our - 8
out - 4
over - 8
own - 8

part - 8
point
put - 7

right - 7

said - 4
same - 8
saw - 5
say - 7
second
see - 3
set
she - 4
should
side
small - 7
so - 1
some - 7
something
still - 8
such

take - 8
than
that - 1
the - 2
their - 8
them - 7
then - 7
there - 7
these
they - 4
thing
think - 8
this - 8
those
thought - 7
through
time - 7
to - 2
told - 7
too
took - 8
toward

under
until
up - 4
us - 8
use

very - 7

want - 5
was - 5
way - 8
we - 7
well
went - 5
were - 7
what - 5
when - 7
where - 6
which
while
who - 5
why - 8
will - 5
with - 5
without - 8
work - 5
would - 8

yet - 7
you - 6
young
your - 7

* variation introduced early

Appendix B Alphabetized Word Lists*

Second 100 Words *(Level 2)*

5- again (139)
almost (160)
along (189)
always (155)
8- another (123)
around (140)
away (156)

8- because (103)
8- began (197)
8- best (191)
better (165)
between (118)
3- big (186)
both (119)

5- came (132)

5- didn't (171)
different (198)
does (151)
done (194)
7- don't (150)
4- down (101)

each (104)
early (184)
end (169)
enough (161)
8- ever (192)
8- every (149)

8- far (162)
felt (188)
few (136)

4- find (172)
8- found (142)

give (178)
4- good (110)
7- got (152)
great (129)

8- help (199)
4- here (117)
high (148)
himself (135)
8- how (107)

8- just (105)

kind (196)
knew (175)
7- know (124)

large (185)
8- last (126)
later (174)
left (153)
7- let (181)
3- little (109)
6- long (116)
3- look (173)

4- make (112)
matter (200)
might (127)

need (187)
8- never (121)
8- next (177)
nothing (168)

8- off (131)
7- old (130)
once (147)
7- open (195)
8- own (114)

8- part (146)
point (176)
7- put (158)

7- right (133)

6- same (122)
5- saw (190)
7- say (145)
second (183)
set (166)
should (102)
side (182)
7- small (141)
something (157)
8- still (113)

6- take (134)
thing (193)
6- think (159)
those (106)
7- thought (143)
7- told (167)
too (108)
6- took (163)
toward (179)

under (120)
until (154)

8- us (128)
use (137)

7- very (111)

5- went (144)
while (125)
8- why (170)
8- without (138)
6- work (115)

7- yet (164)
young (180)

*rank from Modified Frequency List indicated in parentheses

All 200 Words

a
about
after
again
all
almost
along
also
always
an
and
another
any
are
around
as
at
away

back
be
because
been
before
began
best
better
between
big
both
but
by

came
can

come
could

did
didn't
different
do
does
done
don't
down

each
early
end
enough
even
ever
every

far
felt
few
find
first
for
found
from

get
give
go
good
got
great

had
has

have
he
help
her
here
high
him
himself
his
how

I
if
in
into
is
it
its

just

kind
knew
know

large
last
later
left
let
like
little
long
look

made
make

many
matter
may
me
might
more
most
much
must
my

need
never
new
next
no
not
nothing
now

of
off
old
on
once
only
open
or
other
our
out
over
own

part
point
put

right

said
same
saw
say
second
see
set
she
should
side
small
so
some
something
still
such

take
than
that
the
their
them
then
there
these
they
thing
think
this
those
thought
through
time
to

told
too
took
toward

under
until
up
us
use

very

want
was
way
we
well
went
were
what
when
where
which
while
who
why
will
with
without
work
would

yet
you
young
your